Dell EqualLogic Configuration Guide

Dell Storage Engineering

- **Configure unified block and file storage solutions based on EqualLogic PS Series arrays and the FS Series Family of NAS Appliances.**
- **Recommendations and best practices for iSCSI SAN and scale-out NAS network fabric design.**
- **Updated capacity guidelines, capabilities, limitations, and feature sets for the EqualLogic product family.**

Version 14.1 March 2013

Abstract

This configuration guide provides technical guidance for designing and implementing Dell EqualLogic PS Series storage solutions in iSCSI SAN environments. The goal of this guide is to provide a single reference for product technical information and recommended SAN and NAS design methods:

- Details on the capacity and capabilities of different EqualLogic PS Series family and FS Series family of NAS Appliances
- An overview of the peer storage feature set provided by the EqualLogic storage controllers and the NAS feature set provided by the FS Series NAS Service
- iSCSI SAN design guidelines for optimizing redundancy, availability and performance of EqualLogic storage arrays.
- Host configuration guidelines and recommendations for Ethernet switch configuration

This document should be used strictly as a guide in planning an EqualLogic SAN solution. It should not be considered a statement of support for any specific configuration. Actual viability of any configuration will depend on the capabilities of the individual components (switches, network interface controllers, initiators, etc.) that make up the SAN infrastructure.

This configuration guide is intended for storage administrators, SAN designers, storage consultants, or anyone who is considering purchasing or has purchased EqualLogic PS Series Arrays for use in a production storage area network.

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2013 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the DELL logo, and the DELL badge, PowerConnect™, Dell™ EqualLogic™, PowerEdge™ and PowerVault™ are trademarks of Dell Inc. Broadcom™ is a registered trademark of Broadcom Corporation. Cisco® is a registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and other countries. Intel™ is a registered trademark of Intel Corporation in the U.S. and other countries. Microsoft™, Windows™, Windows Server™, and Active Directory™ are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.

Revision history

Revision	Date	Changes
14.1	March 2013	Added section 14 Data Center Bridging Added volume limits for Synchronous Replication in Table 3 Modified section 7.1.2 and section 7.2 Changed all references in tables in section 4 to read PS41x0 and PS61x0 Added note in section 5.1 Updated Appendix D – EqualLogic Upgrade Paths Updated the Related Publications
13.4	December 2012	Updates: New section numbering Updated Preparing for Firmware Upgrade and Controller Failover Added note box to reflect 10Gb support only on PS4110 and PS6110 ports Updated for PS-M4110XS and PS65x0 Hybrid Arrays Added note box to reflect no support for Direct Attach Storage (DAS) New Sections: Appendix D: Upgrade Paths for EqualLogic PS Series Arrays
13.3	September 2012	Updates: Updated for PS-M4110 Array. Updated Blade Integration section to include Dell Force10 MXL. Updated for FS7600/FS7610 and Fluid FS 2.0. Updated Controller Firmware Upgrade Path to Firmware 6.x. Updated Section 2.4 RAID Policies Corrected table 25 in Section 4.7.1. New sections: Section 4 Capacity Planning. Section 6.2 Synchronous Replicaton (SyncRep). Section 6.3 Internet Protocol Security (IPsec).
13.2	June 2012	Added PS Series Firmware Compatibility with EqualLogic Tools in Table 7.
13.1	March 2012	Updated for PS6110/4110. Updated Blade Integration in section 4.7. Updated the capacity tables from raw storage sizes to usable storage sizes. Added replication partner compatibility information.
12.4	November 2011	New sections and other changes throughout the document to include coverage of FS Series NAS Appliance (FS7500)
12.3	October 2011	Updates for PS4100/PS6100 family arrays: controller details; vertical port failover; SAN configuration guidelines; RAID level capacity tables

12.2.1	July 2011	Corrected Spanning Tree portfast guidance in Appendix D.
12.2	May 2011	Updated figures and content additions to Replication section
12.1	March 2011	Section 3.3.7: Replication Section 4.6: Integrating 1GbE and 10GbE SANs Section 2.1: Supported Configuration Limits (maintaining consistency with current firmware release notes) Appendix B and Appendix C: Content removed and now published as a separate **Validated Components List** document

All significant changes included in the most recent version of this document are highlighted using blue text.

Table of contents

Abstract i
Revision history ii
1 PS Series storage arrays 1-2
1.1 Array models 1-2
1.2 PS Series supported configuration limits 1-3
1.3 Array models prior to PS4100/PS6100 1-5
1.3.1 Controller types in all models prior to PS4100/PS6100 1-6
1.3.2 Controller redundancy in all models prior to PS4100/PS6100 1-7
1.3.3 Controller failover behavior in all models prior to PS4100/PS6100 1-7
1.4 Array models PS4100/PS6100 1-10
1.4.1 Controller types in PS4100/PS6100 models 1-10
1.4.2 Controller redundancy in PS4100/PS6100 controllers 1-11
1.4.3 Controller failover behavior: PS41x0/PS61x0 1-11
1.4.4 Vertical port failover behavior in PS4100/PS6100 controllers 1-12
1.4.5 Vertical port failover behavior in PS4110/PS6110 controllers 1-15
1.5 Array model PS-M4110 1-17
1.5.1 Controller type in PS-M4110 model 1-17
1.5.2 Configuration options 1-18
1.5.3 Failure behavior in the PS-M4110 1-18
1.5.4 Networking considerations and guidelines 1-18
2 Controller firmware 2-20
2.1 About member firmware 2-20
2.2 Firmware upgrade considerations 2-20
2.2.1 PS Series Firmware Compatibility with EqualLogic Tools 2-22
2.3 Optimizing for High Availability and preparing for Array Firmware updates 2-23
2.3.1 Verify hosts are designed for high availability 2-23
2.3.2 Microsoft Windows Servers 2-23
2.3.3 VMWare ESX Servers 2-23
2.3.4 Storage Heartbeat on vSphere 5.0, 4.1, and 4.0 2-24
2.3.5 Linux servers 2-24
2.3.6 Network topology health check 2-24
2.3.7 Environmental health 2-24
3 RAID policies 3-25

3.1 Setting the RAID Policy for a member 3-25
3.2 Guidelines for choosing a RAID policy 3-25
3.2.1 RAID level 3-26
3.2.2 Drive configuration 3-26
3.2.3 Spare drive policy 3-26
3.3 RAID level characteristics 3-26
3.3.1 Performance and availability characteristics of the supported RAID levels 3-27
3.4 Supported RAID policy conversions 3-28
4 Capacity planning 4-30
4.1 RAID 6 drive layouts and total reported usable storage 4-30
4.2 RAID 10 drive layouts and total reported usable storage 4-32
4.3 RAID 50 drive layouts and total reported usable storage 4-33
4.4 RAID 5 drive layouts and total reported usable storage 4-34
4.5 Array RAID configurations and associated RAID sets 4-35
5 PS Series array concepts 5-36
5.1 Groups and pools 5-36
5.1.1 Pools 5-36
5.2 Volumes 5-38
5.2.1 Volume attributes 5-38
5.2.2 Volume features 5-39
5.3 Snapshots and clones 5-39
5.3.1 Clones 5-40
5.4 Thin provisioning 5-40
5.4.1 Template volumes and thin clones 5-41
6 Array firmware features 6-42
6.1 Replication 6-42
6.1.1 Replication limits 6-42
6.1.2 Replication paths 6-43
6.1.3 Replication process 6-44
6.1.4 Fast failback 6-46
6.1.5 Sizing Replica Reserve and Delegated Space 6-46
6.1.6 Effect of TCP Window Size on Latency across WAN links 6-48
6.1.7 Replication partner compatibility 6-48
6.1.8 Clustering 6-49

6.2 Synchronous replication 6-49
6.2.1 About Synchronous replication 6-49
6.2.2 How Synchronous replication works 6-49
6.2.3 Synchronous states 6-50
6.2.4 Caveats: 6-50
6.2.5 Requirements for using SyncRep 6-50
6.2.6 How SyncRep protects volume availability 6-51
6.2.7 SyncAlternate volume unavailable 6-51
6.2.8 Tracked changes written to the SyncAlternate volume 6-52
6.2.9 SyncActive volume unavailable 6-53
6.3 Protecting your EqualLogic group with Internet Protocol Security 6-56
6.3.1 Types of protected traffic 6-56
6.3.2 Protected Intra-Group Traffic 6-57
6.3.3 IPsec and Replication 6-57
6.3.4 About IPsec Security Parameters 6-57
6.3.5 About IPsec Certificates 6-57
6.3.6 About IPsec Pre-Shared Keys 6-58
6.3.7 About IPsec policies 6-58
6.3.8 IPsec considerations and limitations 6-58
6.3.9 Performance considerations 6-59
6.3.10 Host Connectivity Considerations 6-59
7 EqualLogic SAN design 7-60
7.1 General requirements 7-60
7.1.1 Implementation of standards 7-60
7.1.2 General requirements and recommendations 7-61
7.1.3 Quality of service (qos) 7-62
7.2 Ethernet switches and infrastructure 7-62
7.2.1 Connecting SAN switches in a Layer 2 network 7-64
7.2.2 Sizing inter-switch connections 7-67
7.2.3 Comparing inter-switch connection types 7-67
7.3 Building a high-availability SAN 7-68
7.3.1 Design guidelines for host connectivity in a redundant SAN 7-69
7.3.2 Multi-path I/O 7-70
7.3.3 Equallogic iSCI SAN Design 7-71

7.3.4 Redundant SAN configuration 7-72
7.3.5 Partially redundant SAN configurations 7-74
7.3.6 Minimum cabling scenarios: PS4100 and PS6100 7-75
8 Mixed speed environments - Integrating 1GbE and 10GbE SANs 8-78
8.1 Design considerations 8-78
8.1.1 Optimizing Rapid Spanning Tree Protocol behavior 8-79
8.2 Mixed speed SAN best practices 8-80
9 Blade server chassis integration 9-82
9.1 Designing a SAN using blade chassis I/O modules with arrays directly attached 9-84
9.1.1 SAN design for multiple M1000e enclosure 9-90
9.2 Designing a SAN using blade Pass-through module 9-94
9.3 Designing a SAN using blade chassis I/O modules as host access to external switches for array connection 9-96
10 Fluid File system 10-103
10.1 FS Series architecture 10-104
10.1.1 FS Series solution for file only storage 10-104
10.2 Dell FluidFS 10-105
11 FS Series NAS Appliances 11-107
11.1 Equallogic NAS appliance supported configuration limits 11-107
11.2 Initial NAS cluster valid configurations 11-109
11.3 FS7500 system components 11-110
11.4 FS7500 file system operation on controller failover 11-111
11.5 FS7600 components 11-111
11.6 FS7610 components 11-112
12 FS Series file level operations 12-113
12.1 NAS cluster 12-113
12.2 NAS reserve 12-113
12.2.1 Relationship between PS Series groups, pools and NAS reserve 12-113
12.3 NAS Container 12-114
12.3.1 NAS Container security styles 12-115
12.4 NAS Container snapshots 12-116
12.5 NAS Snapshots and Replication 12-117
12.5.1 Snapshots 12-117
12.5.2 Replication 12-117
13 FS Series NAS Configuration 13-120

13.1 FS7500 connection paths 13-120
13.2 FS7600/7610 connection paths 13-122
14 Data Center Bridging (DCB) 14-126
14.1 DCB Overview 14-126
14.2 DCB requirements for EqualLogic 14-129
14.3 Methods for configuring DCB 14-130
14.4 Basic Deployment Topology Example 14-131
14.5 Tested SAN designs 14-131
14.5.1 Blade IOM switch only 14-132
14.5.2 ToR switch only 14-132
14.5.3 Blade IOM switch with ToR switch 14-133
14.6 Data Center In A Chassis DCB design 14-134
14.7 VLANs for iSCSI 14-135
Appendix A Network ports and protocols 14-137
A.1 Required ports and protocols 14-137
A.2 Optional ports and protocols 14-137
Appendix B Recommended switches 14-139
Appendix C Supported iSCSI initiators 14-140
Appendix D Upgrade paths for EqualLogic PS SeriesArrays 14-141
Related publications 14-142

Introduction

With the Dell™ EqualLogic™ PS Series of storage arrays Dell provides a storage solution that delivers the benefits of consolidated networked storage in a self-managing, iSCSI storage area network (SAN) that is affordable and easy to use, regardless of scale. By eliminating complex tasks and enabling fast and flexible storage provisioning, these solutions dramatically reduce the costs of storage acquisition and ongoing operations.

To be able to leverage the advanced features provided by an EqualLogic array, a robust, standards-compliant iSCSI storage area network (SAN) infrastructure must be created. While almost any industry standard, enterprise class Ethernet switch can be used to connect hosts to the EqualLogic-based virtual storage volumes, some switches, network interface controllers(NICs), iSCSI Host Bus Adapters (HBAs), and iSCSI initiators can directly impact the overall performance of the SAN storage solution. The *Dell EqualLogic Configuration Guide* is provided as an aid to help storage administrators determine how best to build an iSCSI infrastructure for use within an EqualLogic SAN solution. This document focuses on network configuration, host integration, and other topics that help to ensure a smooth deployment with optimum PS Series SAN and FS Series NAS appliance performance.

> **Note:** This version of the EqualLogic Configuration Guide (v13.3, September 2012) introduces coverage of the EqualLogic PS-M4110 Blade Array. Unless otherwise stated, the recommendations involving specific 10GbE solutions apply to the first generation PS6010, the second generation PS4110, PS-M4110, and the PS6110.

Unless otherwise stated, recommendations in this document are applicable to both file and block environments.

This document should be used strictly as a guide in planning an EqualLogic storage solution. It should not be considered a statement of support for any specific configuration. Actual viability of any configuration will depend on the capabilities of the individual components (switches, network interface controllers, initiators, etc.) that make up the SAN infrastructure.

Audience

This configuration guide is intended for storage administrators, SAN/NAS system designers, storage consultants, or anyone who is considering purchasing or has purchased EqualLogic PS Series storage or FS Series Array appliances for use in a production storage area network. It is assumed that all readers have experience in designing and/or administering a shared storage solution. Also, there are some assumptions made in terms of familiarity with all current and possibly future Ethernet standards as defined by the Institute of Electrical and Electronic Engineers (IEEE) as well as all TCP/IP and iSCSI standards as defined by the Internet Engineering Task Force (IETF).

1 PS Series storage arrays

PS Series Storage SANs provide a peer storage architecture comprised of one or more independent arrays. Each array contains its own controllers, cache, storage, and interface ports. Grouped together they can create one or more single instance storage pools that are based on the IETF iSCSI standard. The capabilities of the PS Series solution are provided through a unique set of software and hardware capabilities that, when used cooperatively, can provide a full featured SAN solution. The following sections provide additional insight into specifications and functions of PS Series arrays.

1.1 Array models

The array models available prior to introduction of the PS4100/PS6100 family are shown in Table 1 below.

Table 1 Array Models prior to PS4100/PS6100

Array Model	Drive Type	Number of Drives
PS50E – PS2400E	SATA	14, 16[(a)]
PS3000X, PS3x00XV	SAS	16
PS4000E	SATA	16
PS4000X, PS4000XV	SAS	16
PS5000E	SATA	16
PS5000X, PS5000XV	SAS	16
PS5500E	SATA	48
PS6000E	SATA	16
PS6000X, PS6000XV	SAS	16
PS6000S, PS6010S	SSD	16
PS6010E	SATA	16
PS6010X, PS6010XV	SAS	16
PS6000XVS, PS6010XVS	SAS / SSD	8 / 8
PS6500E	SATA	48
PS6510E	SATA	48
PS6510X	SAS	48
PS6500ES	SAS / SSD	41 SAS + 7 SSD
PS6510ES	SAS / SSD	41 SAS + 7 SSD

(a) PS2400E comprised of four drive enclosures with 14 drives each

Starting with the introduction of the PS4100 and PS6100 family of arrays, configurations using 2.5" and 3.5" disks are available.

Table 2 PS4100/PS6100 array models

Array model	Drive type	Number of drives
PS4100E	3.5" SAS 7.2K RPM	12
PS4100X	2.5" SAS 10K RPM	24
PS4100XV	2.5" SAS 15K RPM	24
PS4100XV	3.5" SAS 15K RPM	12
PS6100E	3.5" SAS 7.2K RPM	24
PS6100X	2.5" SAS 10K RPM	24
PS6100XV	2.5" SAS 15K RPM	24
PS6100XV	3.5" SAS 15K RPM	24
PS6100S	SSD	12 or 24
PS6100XS	SSD + SAS 10K RPM	7 SSD + 17 SAS
PS4110E	3.5" SAS 7.2K RPM	12
PS4110X	2.5" SAS 10K RPM	24
PS4110XV	2.5" SAS 15K RPM	24
PS4110XV	3.5" SAS 15K RPM	12
PS6110E	3.5" SAS 7.2K RPM	24
PS6110X	2.5" SAS 10K RPM	24
PS6110XV	2.5" SAS 15K RPM	24
PS6110XV	3.5" SAS 15K RPM	24
PS6110S	SSD	12 or 24
PS6110XS	SSD + SAS 10K RPM	7 SSD + 17 SAS
PS-M4110E	7.2K RPM NL-SAS 2.5"	14
PS-M4110X	10K RPM SAS 2.5"	14
PS-M4110XV	15K RPM SAS 2.5"	14
PS-M4110XS	SSD + 10K RPM 2.5" SAS	5 SSD + 9 SAS

1.2 PS Series supported configuration limits

The Dell EqualLogic PS6xxx Series provides the full range of features and capabilities available with the EqualLogic PS Series storage products. The Dell EqualLogic PS4xxx Series provides a subset of features and capabilities, targeted at remote office and small to medium business storage deployments. The supported configuration limits for a PS Series group provided in Table 3.

Table 3 Supported configuration limits

Configuration	PS4000/PS4100 and PS-M4110 groups only[a]	All other groups[b]
Volumes and replica sets per group	512[h]	1024
Volume size[c]	15 TB	15 TB
Volumes enabled for replication (outbound)[d]	32	256
Snapshots and replicas per group	2048	10,000
Snapshots per volume	128	512
Replicas per volume	128	512
Volumes that have Synchronous Replication Enabled	4	32
Schedules (snapshot or replication) per volume or volume collection	16	16
Persistent Reservation registrants per volume	96	96
Replication partners per group	16	16
Replication partners per volume	1	1
Members per group	2	16[a]
Members per pool	2	8
Pools per group	2	4
Volumes per collection	8	8
Collections per group (snapshot and replication)	100	100
Volume connections (each time an iSCSI initiator connects to a volume counts as a connection)[e,f]	512 per pool[g] 1024 per group with 2 pools	1024 per pool[h] 4096 per group with 4 pools
Access control records per volume and its snapshots	16	16
Simultaneous management sessions (any combination of GUI, telnet, or scripting sessions)	7	7
Thin provisioning[i] limits (minimum allocation)	10% of volume size	10% of volume size
Administrator accounts per group	100	100

(a) A group can contain a maximum of two PS4000, PS4100, and/or PS-M4110 arrays.
(b) Includes groups of mixed array types and all other group types except groups containing only PS40x0, PS41x0, and PS-M4110 arrays. When a group contains a mix of PS40x0, PS41x0, and PS-M4110 arrays and other array models, the higher limits prevail.
(c) Practical maximum volume size is operating system-specific. A PS Series group can create and present volumes up to 15 TB.
(d) Up to 16 simultaneous replication operations can take place.
(e) To avoid service outages or possible loss of connectivity during failovers, Dell recommends increasing initiator timeouts for groups with more than 512 connections. See the iSCSI Initiator and Operating Systems Considerations document for more information.
(f) Inbound replication connections count towards the total number of connections.
(g) With firmware version 5.1, maximum number of volumes per group and maximum number of connections per pool increased from 256 to 512
(h) Up to 1024 connections per pool are supported for all groups except groups containing only PS40x0, PS41x0, and PS-M4110 arrays. This provides increased scalability in virtualized environments. These connections may be distributed across a maximum of 512 volumes.
(i) Volumes automatically created and managed by the NAS Service inside the NAS Reserve pool used by an FS Series appliance are fully allocated at creation and cannot be thin provisioned. Once those volumes are created (when the NAS Reserve is created they cannot be shrunk.

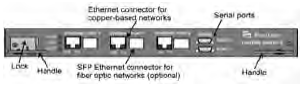
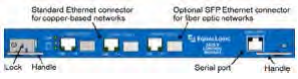
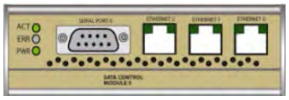
1.3 Array models prior to PS4100/PS6100

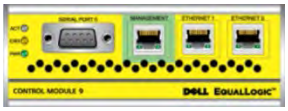
Since the EqualLogic PS Series was introduced, there have been several different array models released with new features, better performance and greater storage capacity. The storage array controllers were also improved to take advantage of advances in the underlying networking and storage technologies.

1.3.1 Controller types in all models prior to PS4100/PS6100

Array controllers can be identified and differentiated by the controller "type" designation. Each controller type will have a different colored label to help quickly identify the controller type. Table 4 lists each Dell EqualLogic controller along with some characteristics about each.

Table 4 Array controller types – All models prior to PS4100/PS6100

Controller type	Faceplate	Network interfaces	Storage type	Notes
Type 1		3 x 1GbaseT 3 x 1Gb SFP (combo)	SATA	Original Controller Design PS50 – PS2400 1GB Cache
Type 2		3 x 1GbaseT 3 x 1Gb SFP (combo)	SATA	PS50 – PS2400 1GB Cache
Type 3 SAS Type 3 SATA		3 x 1GbaseT	SAS SATA	PS3000 – PS5000 1GB Cache Cannot mix Type 3 SAS with Type 3 SATA
Type 4		3 x 1GbaseT	SAS	PS3000 – PS5000 1GB Cache Cannot mix Type 3 SAS with Type 4 controller
Type 5		3 x 1GbaseT	SATA	PS3000 – PS5000 1GB Cache Cannot mix Type 3 SAS with Type 5 controller
Type 6		3 x 1GbaseT	SATA	PS5500 only 2GB Cache
Type 7		4 x 1GbaseT	SAS SATA SSD	PS6000 – PS6500 2GB Cache Cannot mix SAS and SATA drives in one array
Type 8		2 x1GbaseT 1 x 10/100Mb mgt	SAS SATA	PS4000 only 2GB Cache Cannot mix SAS and SATA drives in same array

Controller type	Faceplate	Network interfaces	Storage type	Notes
Type 9		2 x 1GbaseT 1 x 10/100Mb mgt	SAS SATA	2nd generation PS4000 2GB Cache Cannot mix SAS and SATA drives in same array
Type 10		2 x 10GB SFP+ 1 x 10/100Mb mgt	SAS SATA SSD	10Gb Ethernet PS6010 – PS6510 2GB Cache

1.3.2 Controller redundancy in all models prior to PS4100/PS6100

Each array can be configured with either a single controller, or dual redundant controllers. The single controller configuration will provide the same level of I/O performance as a dual controller configuration. The dual controller configuration provides for redundancy. Redundant controllers will prevent volume connections between hosts and SAN from being dropped in the event of an active controller failure.

The *Active Controller* is the controller which is processing all disk and network I/O operations for the array. A second controller in dual controller configurations will always be in a "passive" operating mode. In this mode, the secondary controller will exhibit the following characteristics:

- Each of its Ethernet ports are electrically inactive (active lights are off)
- The passive controller cache mirrors the cache of the active controller.

1.3.3 Controller failover behavior in all models prior to PS4100/PS6100

To support redundant controller fail over, each Ethernet port on the active controller that is connected to the SAN must have its corresponding port on the passive controller also connected to the same SAN network. In the event of a controller failure, the passive controller will immediately activate and continue to process all data requests to the array. The following changes occur during fail over:

- The backup controller automatically enables each of the corresponding Ethernet ports that were enabled on the failed primary controller
- The IP addresses that were assigned to each of the failed controller Ethernet ports are reassigned to the corresponding ports on the second controller.

A link failure on one or more of the Ethernet ports on the active controller does not constitute a controller failure. For this reason, it is important to ensure that ports from each controller are connected to at least two different switches. This will prevent a switch failure from also disabling all paths between a host and its connected volumes.

It is also critical that port connections for both controllers are configured so that the corresponding ports on each controller are connected to the SAN. If port 0 on the active controller is the only port

connected to the SAN, then you must use port 0 on the passive controller for the other connection to the SAN. This is illustrated in the partially connected scenario shown in Figure 1.

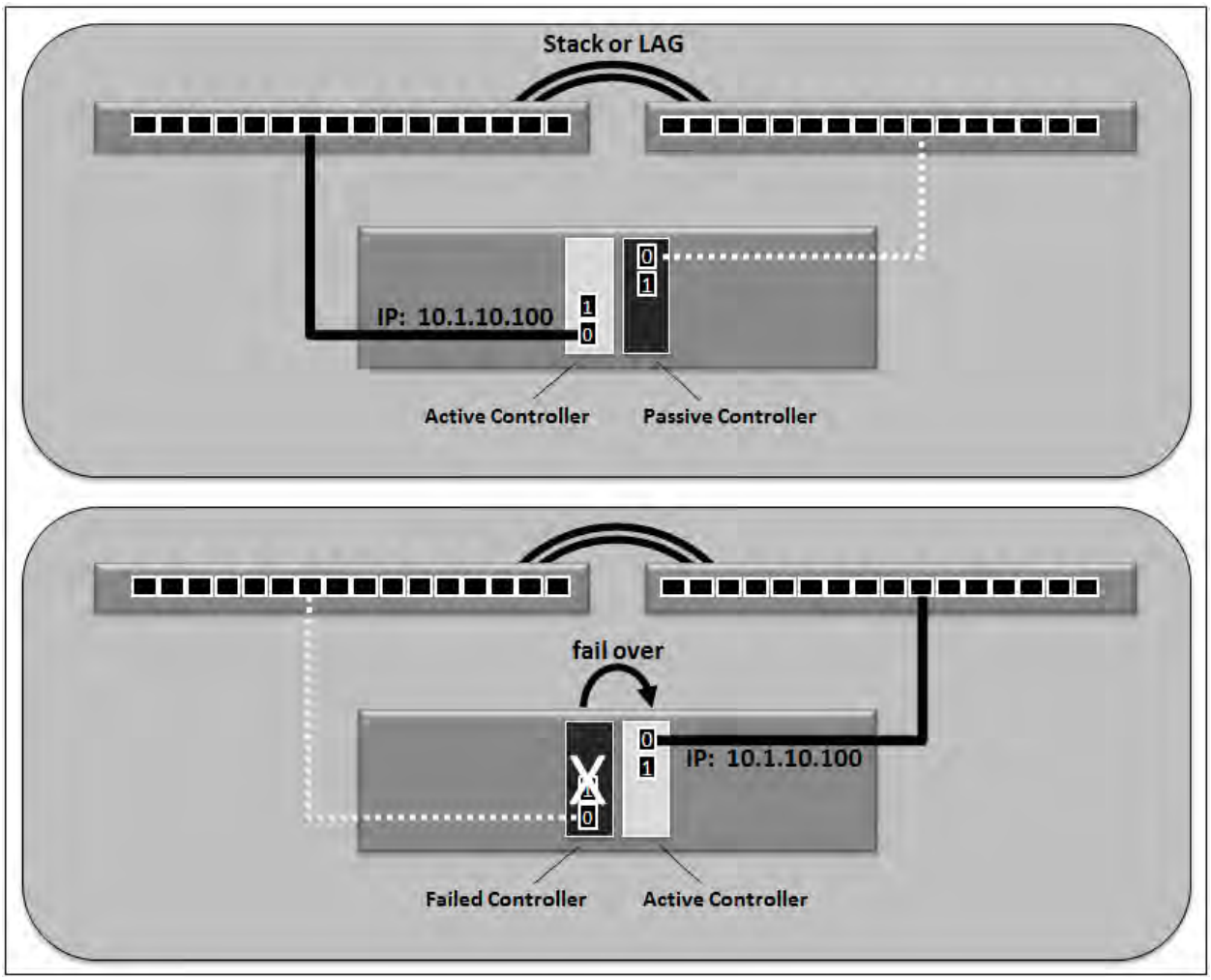


Figure 1 Partially connected controller failover

Note how IP addresses are reassigned to the ports during the failover processes shown in Figure 1 and Figure 2.

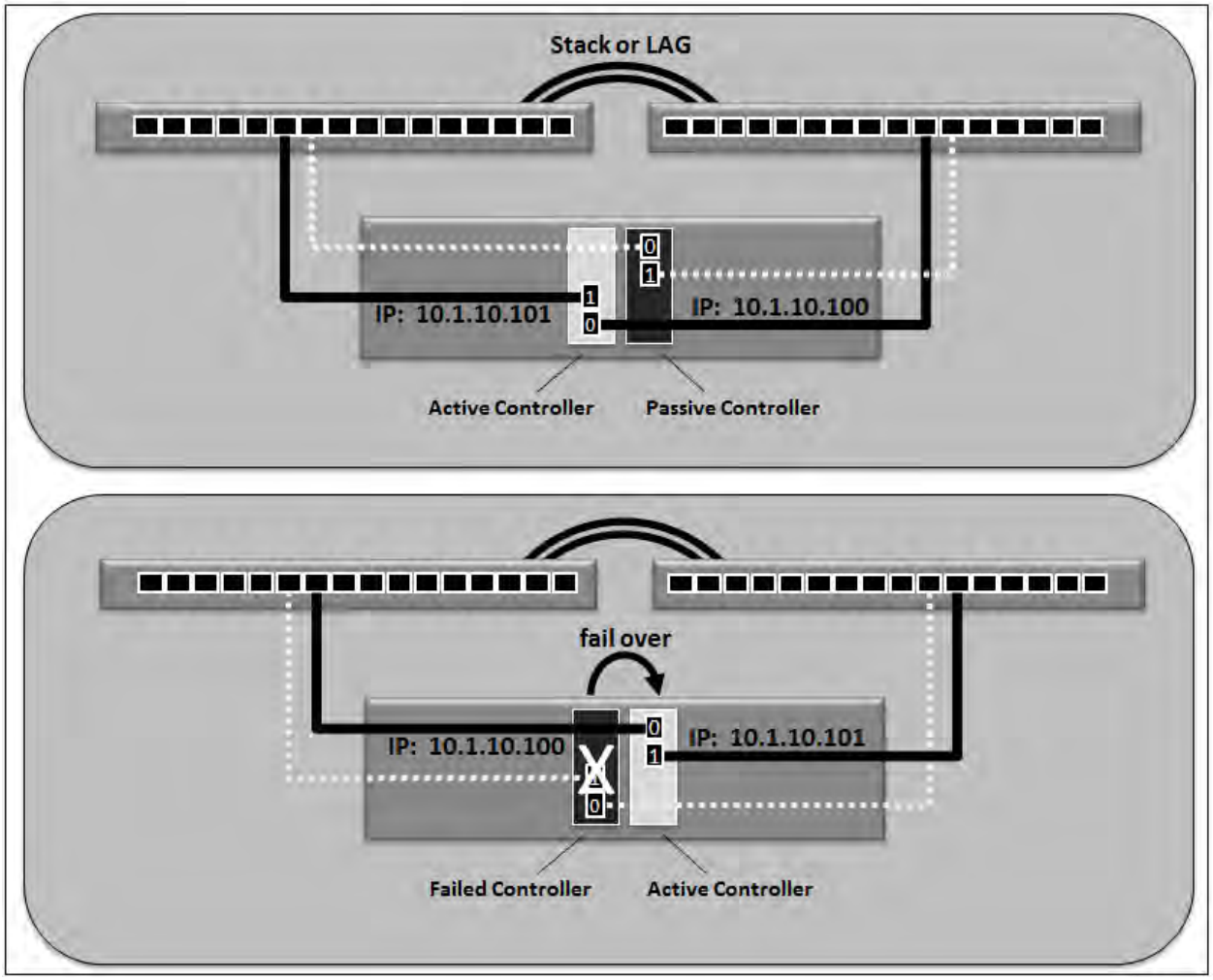


Figure 2 Fully connected controller failover

1.4 Array models PS4100/PS6100

1.4.1 Controller types in PS4100/PS6100 models

The new controller types available in the PS4100 and PS6100 model arrays became available starting in August 2011. Table 5 lists each Dell EqualLogic controller along with some characteristics.

Table 5 PS4100/PS6100 controller types

Controller Type	Faceplate	Network Interfaces	Storage Type	Notes
Type 11		4 x 1GbaseT 1 x 10/100Mb mgt	SAS NL-SAS SSD	PS6100 Only 4GB Cache Supports vertical port failover Cache to Flash memory destaging
Type 12		2 x 1GbaseT 1 x 10/100Mb mgt	SAS NL-SAS	PS4100 Only 4GB Cache Supports vertical port failover Cache to Flash memory destaging
Type 14		1 x 10GbE SFP+ 1 x 10GbaseT 10/100Mb mgt	SAS NL-SAS SSD	PS6110 Only 4GB Cache Support for vertical port failover Cache to Flash memory destaging SFP+ OR 10GBaseT used at any given time
Type 17		1 x 10GbE SFP+ 1 x 10GbaseT 1 x 10/100Mb mgt	SAS NL-SAS	PS4110 Only 4GB Cache Supports vertical port failover Cache to Flash memory destaging SFP+ OR 10GBaseT used at any given time

1.4.2 Controller redundancy in PS4100/PS6100 controllers

Each array can be configured with either a single controller, or dual redundant controllers. The single controller configuration will provide the same level of I/O performance as a dual controller configuration. The dual controller configuration provides for redundancy. Redundant controllers will prevent volume connections between hosts and SAN from being dropped in the event of an active controller failure.

The *Active Controller* is the controller which is processing all disk and network I/O operations for the array. A second controller in dual controller configurations will always be in a "passive" operating mode. In this mode, the secondary controller will exhibit the following characteristics:

- Each of its Ethernet ports are electrically inactive (active lights are off), unless a vertical port failover has occurred.
- The passive controller cache mirrors the cache of the active controller.

1.4.3 Controller failover behavior: PS41x0/PS61x0

In the event of a controller failure the following changes occur:

- The passive controller immediately activates and continues to process all data requests to the array.
- Vertical port pairing insures that IP addresses assigned to each of the failed controller Ethernet ports apply to the corresponding ports on the second controller.

As stated in Section 1.3.3 above, controller models prior to the PS4100/PS6100 required having cabled connections present on corresponding controller ports for controller failover to work correctly. The vertical port failover feature in PS41x0/PS61x0 controllers eliminates this requirement, enabling partially cabled configurations to support controller failover too. See Section 7.3.6, *Minimum cabling scenarios: PS4100 and PS6100*, for details.

It is important to ensure that ports from each controller are connected to at least two different switches. This will prevent a switch failure from also disabling all paths between a host and its connected volumes. We illustrate controller failover behavior for the PS4100 family controller in Figure 3. Controller failover behavior for the PS6100 (4 port) controller family is identical.

IMPORTANT: To prevent a switch failure from also disabling all paths between a host and its connected volumes, you should make sure that ports from each controller are connected to at least two different switches. **You should also split vertical port pair connections between two switches to ensure 100% bandwidth capability is maintained in the event of a vertical port failover event.** Both of these guidelines are illustrated in Figure 3.

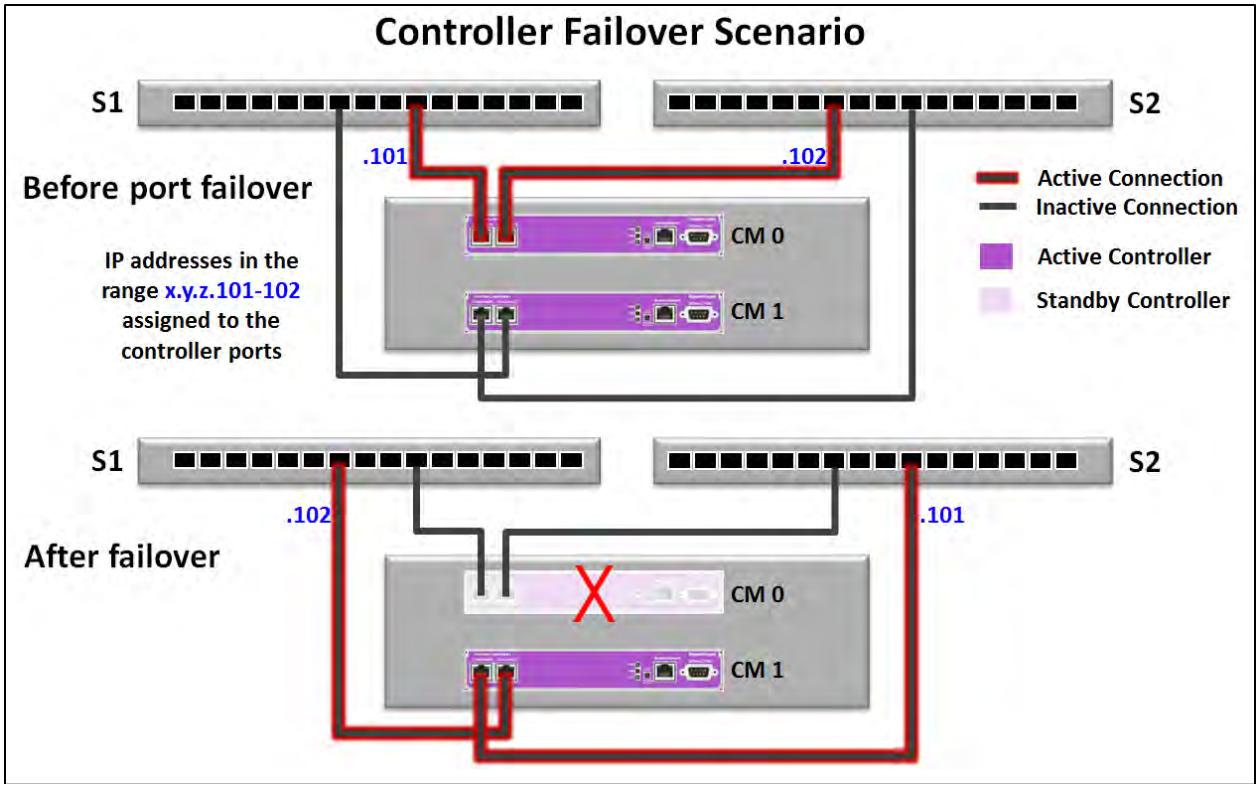


Figure 3 Controller failover process and optimal connection paths

1.4.4 Vertical port failover behavior in PS4100/PS6100 controllers

In PS Series controllers prior to PS4100/6100 families, a link failure or a switch failure was not recognized as a failure mode by the controller. Thus a failure of a link or an entire switch would reduce bandwidth available from the array. Referring to Figure 4 or Figure 5, assume that CM0 is the active controller. In vertical port failover, if CM0 senses a link drop on the local ETH0 port connection path, it will automatically begin using the ETH0 port on the backup controller (CM1). Vertical port failover is bi-directional. If CM1 is the active controller then vertical port failover will occur from CM1 ports to CM0 ports if necessary.

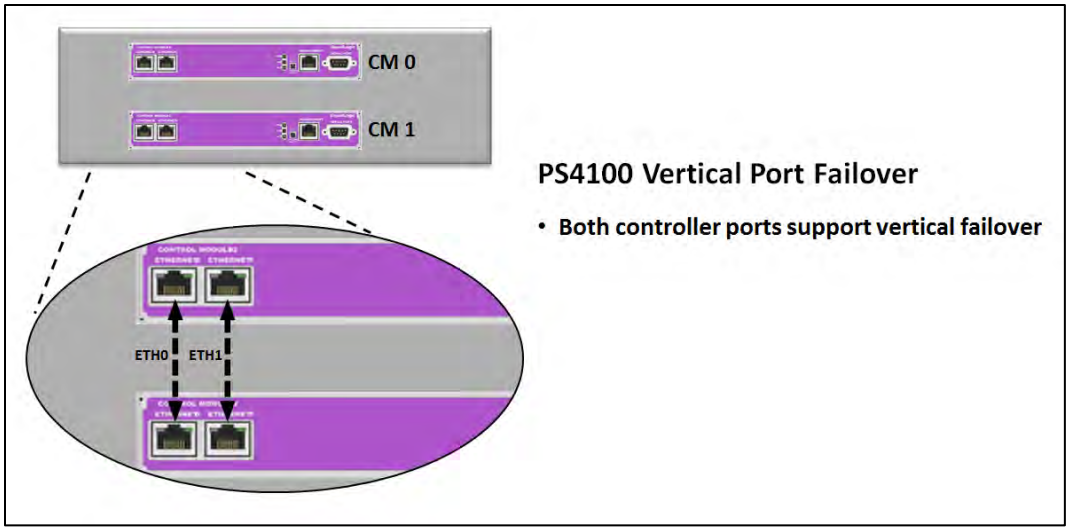


Figure 4 PS4100 vertical port failover

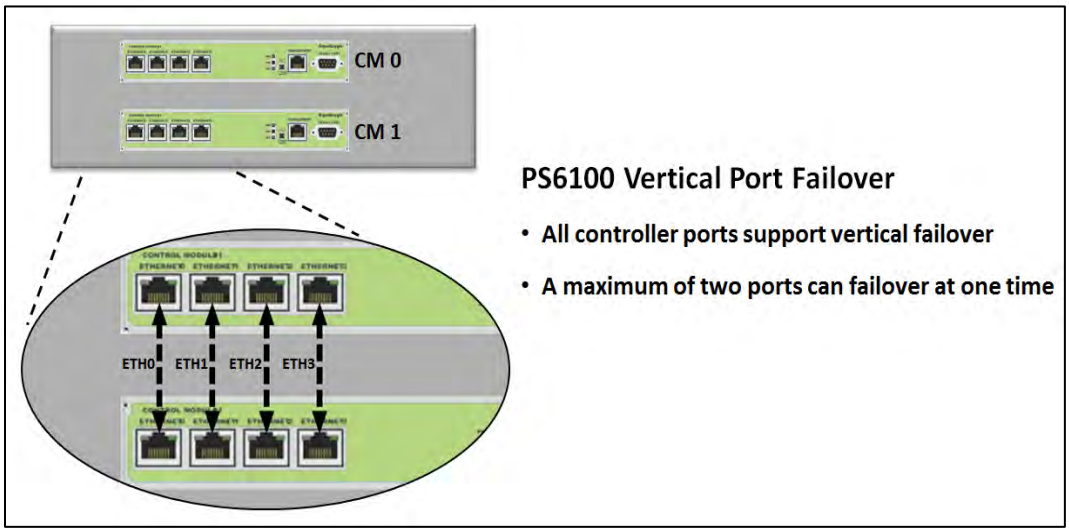


Figure 5 PS6100 vertical port failover

With PS4100/PS6100 family controllers, vertical port failover can ensure continuous full bandwidth is available from the array even if you have a link or switch failure. This is accomplished by combining corresponding physical ports in each controller (vertical pairs) into a single logical port from the point of view of the active controller. In a fully redundant SAN configuration, you must configure the connections as shown in Figure 24 and Figure 25 in section 7.3.4.

In a redundant switch SAN configuration, to optimize the system response in the event you have a vertical port failover **you must split the vertical port pair connections between both SAN switches**. The connection paths illustrated in Figure 6 and Figure 7 show how to alternate the port connection paths between the two controllers. Also note how IP addresses are assigned to vertical port pairs.

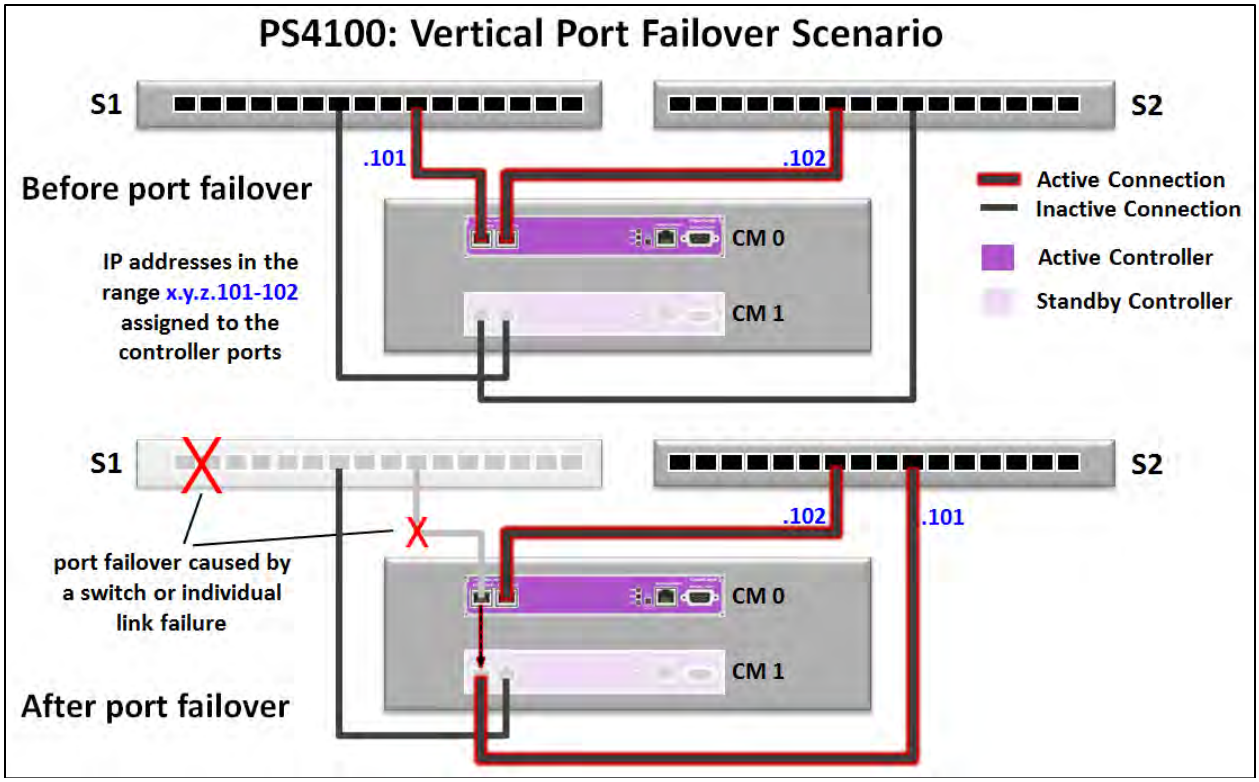


Figure 6 PS4100 vertical port failover and optimal connection paths

IMPORTANT: By alternating switch connection paths between ports in a vertical port pair, port failover allows the array to maintain 100% bandwidth capability in the event of a switch failure.

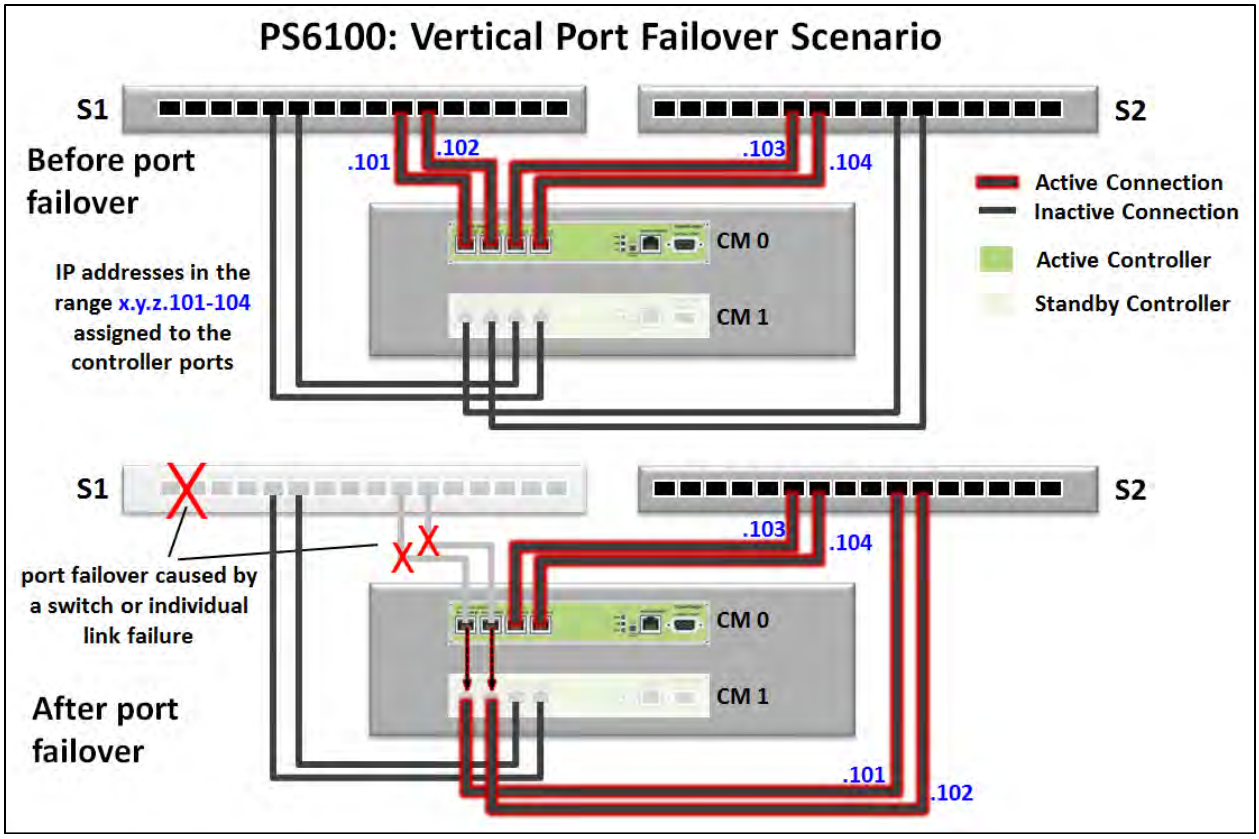


Figure 7 PS6100 vertical port failover process and optimal connection paths

1.4.5 Vertical port failover behavior in PS4110/PS6110 controllers

In PS Series controllers prior to PS4110/6110 families, a link failure or a switch failure was not recognized as a failure mode by the controller. This caused a failure of a link or an entire switch to reduce bandwidth available from the array. Referring to Figure 4 or Figure 5, assume that CM0 is the active controller. In vertical port failover, if CM0 senses a link drop on the local ETH0 port connection path, it will automatically begin using the ETH0 port on the backup controller (CM1). Vertical port failover is bi-directional. If CM1 is the active controller then vertical port failover will occur from CM1 ports to CM0 ports if necessary.

Note: The PS4110 and PS6110 PS Series Arrays do not support dual speed connectivity. They are not designed to support the auto-negotiate feature. The connectivity must be 10Gb ONLY - This is the only supported configuration.

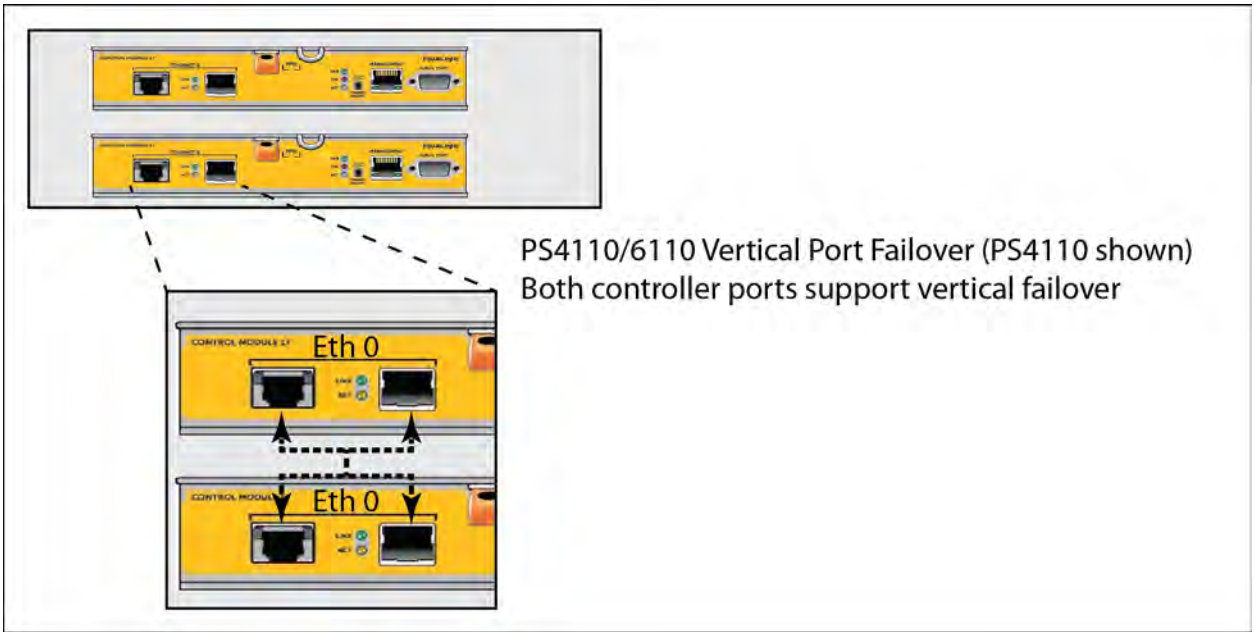


Figure 8 4110/6110 vertical port failover

With the PS4110/PS6110 family of controllers, vertical port failover can ensure continuous full bandwidth is available from the array even if you have a link or switch failure. This is accomplished by combining 10GbE "eth0" ports in each controller into a single logical port from the point of view of the active controller. In a fully redundant SAN configuration, you must configure the connection as shown in Figure 9.

In a redundant switch SAN configuration, to optimize the system response in the event you have a vertical port failover **you must connect either the SFP+ cable or the 10GbaseT cable from each controller to a different switch in the SAN network**. While it is a supported configuration, it is not recommended to connect both the SFP+ and 10GbaseT ports on each controller at the same time. In this scenario, the SFP+ port will always be the preferred active port and this preference cannot be changed by the user. This preference is not guaranteed in future array firmware releases.

The connection paths illustrated in Figure 9 show how to connect the port connection paths between the two controllers. Also note how IP addresses are assigned to the vertical port pair.

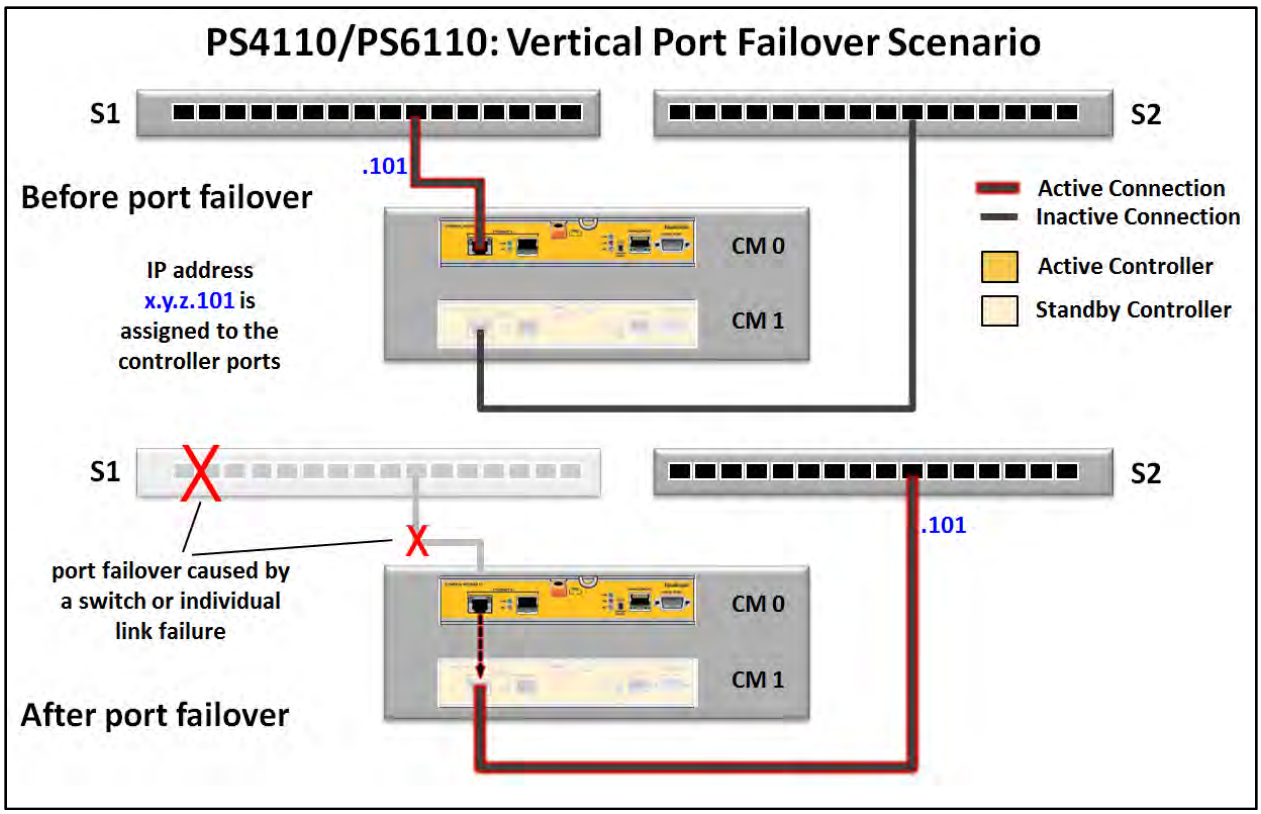


Figure 9 4110/6110 Vertical port failover scenario

1.5 Array model PS-M4110

1.5.1 Controller type in PS-M4110 model

The PS-M4110 controller is designed based on a modified version of the PS4100 Controller. Host and SAS cards combined to form a single unit, fixed I/O module, connecting to the M1000e chassis infrastructure.

Controller Type	Storage blade image	Network Interfaces	Storage type	Notes
Type 13		2 x 10Gb Ports, (One Per Controller), Connected through the Backplane (No Label) Each of the two ports has an active link and a standby link to the fabric switches in the backplane of the M1000e chassis.	SAS NL-SAS	Dual, hot-pluggable 10GbE controllers 4GB of memory per controller, Cache to Flash Design 1 x dedicated management port – accessible via CMC 14x 2.5" 6Gb/s SAS disk drives

1.5.2 Configuration options

The PS-M4110 has four basic configuration options. It can be configured on Fabric A or Fabric B, and each Fabric Configuration can use a 10Gb KR switch or a 10Gb KR Pass-Thru Module (PTM). Figure 10 depicts a basic configuration using MXL switch, however any of the supported switches can be used in this configuration.

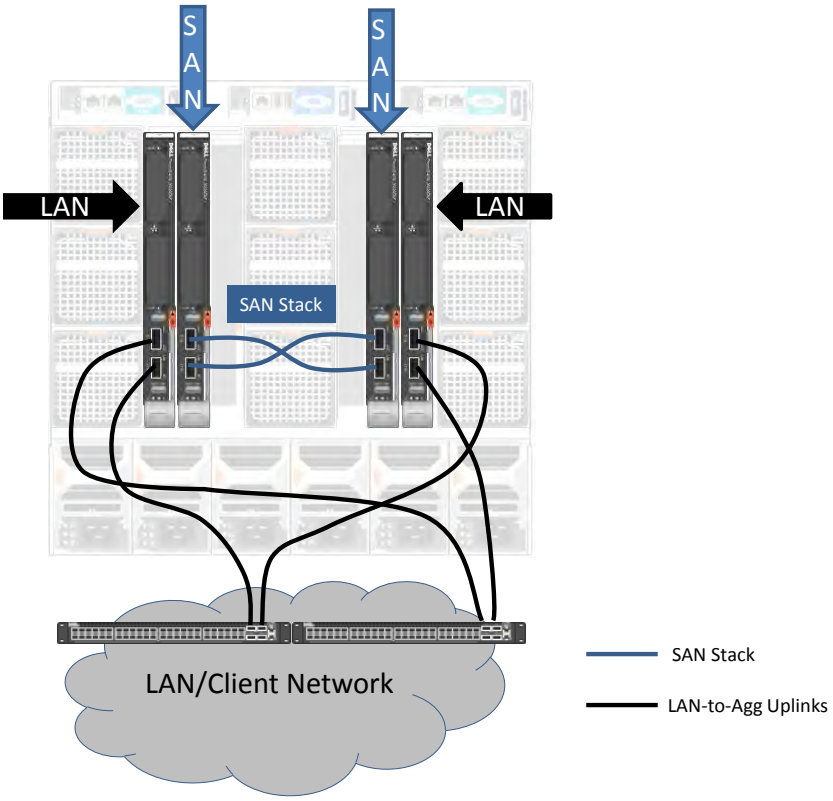


Figure 10 Basic PS-M4110 configuration for data center-in-a-box

1.5.3 Failure behavior in the PS-M4110

Each PS-M4110 array is configured with dual controllers, active and passive connections, and redundant fabrics to provide redundancy and prevent volume connections between hosts and SAN from being dropped in the event of a failure.

There are three failure scenarios that are protected by the array:

- In the event of a link failure, the active link of the active port goes down, the standby link for the same port will take over.
- In the event of a Switch failure, both active and passive ports will automatically link to the redundant fabric switch.
- If the active controller fails, the passive controller will takeover.

1.5.4 Networking considerations and guidelines

Supported M-Series I/O modules

- 10G KR is the only supported IOM
- Switches: Force10 MXL, PowerConnect M8024-k, M8428-k
- Pass-through: 10Gb-K pass-thru modules only with external switches

The following are basic networking recommendations for implementing the PS-M4110 Storage Blade.

- IOMs must be Interconnected
- Datacenter in a Box: IOMs directly interconnected
- External Switches can be used to provide interconnection if rack mounted arrays are needed.
- Must use "Single Fabric" I/O module Placement; Supported Fabrics are "A" and "B" only
- Fabric A not supported on older M1000e systems (Mid-plane v1.0); Mid-plane upgrade kit available
- Fabric B supported on all M1000e system
- It is recommended that PS-M4110 be placed into their own pool

2 Controller firmware

2.1 About member firmware

Each control module in a group member must be running the same version of the PS Series firmware. Firmware is stored on a compact flash card or a microSD card on each control module.

Dell recommends the following:

- Always run the latest firmware to take advantage of new features and fixes.
- All group members must run the same firmware version. If you are adding a new array to a group, update the group to the latest firmware before adding the new member.
- Dell does not support groups running with mixed-firmware versions, except when you are in the process of upgrading the firmware of the entire group, and then only for as long as it takes to do the upgrade.
- You can upgrade member firmware to a higher version or downgrade member firmware to a lower version. In some firmware releases, firmware downgrades are not allowed. See the Dell EqualLogic PS Series Storage Arrays Release Notes for version-specific information about the firmware running on your group member(s).
- For instructions on how to update the firmware, see the Updating Firmware for Dell EqualLogic PS Series Storage Arrays, available at the EqualLogic customer support site (http://www.equallogic.com/support).

Each EqualLogic PS Series array runs a core operating system in firmware that provides all of the PS Series features and functionality. The firmware version is defined using a version number and will be updated from time to time as new features are added or for general maintenance improvements.

The firmware version number takes the following form: "*X.Y.Z*":

- "**X**" is used to identify the "major" release number. This changes when there are very significant changes in the features and functionality.
- "**Y**" is the "minor" release number. This changes as new minor features or improvements in existing features are made for a given major release.
- "**Z**" denotes maintenance releases. This changes whenever a maintenance release level is released to resolve any identified technical or usability issues.

2.2 Firmware upgrade considerations

Before beginning a firmware upgrade process, you must review the Release Notes for the firmware update revision. The release notes will be included with the firmware update kit. All firmware update kits are available at http://www.equallogic.com/support/ (Support ID required for login access). The release notes contain the latest information and guidelines for:

- Completing the upgrade with minimal potential for service disruption
- Operating system and iSCSI initiator configuration settings that will minimize potential for I/O disruption during the upgrade process
- Latest updates to the supported firmware upgrade paths

In addition to the Release Notes, the process for updating controller firmware is described in detail in the following document (Support ID required for login access):

- PS Series Storage Arrays: Updating Storage Array Firmware, available at: https://www.equallogic.com/support/download_file.aspx?id=594
- Supported firmware upgrade paths (up to version 6.0.x) are shown in Table 6 below.
- If you are starting with v4.2.* or v4.3.* then you can update straight to v5.0.4. If the array is already running v5.0 or v5.0.1 then you must first update to v5.0.2 before updating to v5.0.4.

Table 6 Controller firmware upgrade paths

Starting firmware version	Ending firmware version
5.2.x	Later 5.2.x releases 6.0.x
5.1.x	Later 5.1.x releases 5.2.x
5.0.x (excluding 5.0.0 and 5.0.1)	Later 5.0.x releases 5.1.x, 5.2.x
5.0.0, 5.0.1	5.0.2
4.3.x	Later 4.3.x releases 5.0.x (excluding 5.0.0 & 5.0.1)
4.2.x	Later 4.2.x releases 4.3.x 5.0.x (excluding 5.0.0 & 5.0.1)
4.1.x	Later 4.1.x releases 4.2.x, 4.3.x
4.0.x	Later 4.0.x releases 4.1.x
3.3.x	Later 3.3.x releases 4.0.x
3.2.x	Later 3.2.x releases 3.3.x 4.0.x
3.1.x	Later 3.1.x releases 3.2.x, 3.3.x
3.0.x	Later 3.0.x releases 3.1.x

2.2.1 PS Series Firmware Compatibility with EqualLogic Tools

The following table provides a quick reference of EqualLogic product version compatibility for the recent major firmware releases.

Table 7 PS Series firmware compatibility

	Firmware V6.0.x	Firmware V5.2.x	Firmware V5.1.x	Firmware V5.0.x	Firmware V4.3.x
SAN HeadQuarters	2.2.x	2.2.x 2.1.x	2.2.x 2.1.x	2.2.x 2.1.x	2.2.x 2.1.x
Host Integration Tools for Microsoft	4.0.x 3.5.1	4.0.x 3.5.x	4.0.x 3.5.x	4.0.x 3.5.x 3.4.x 3.3.x	4.0.x 3.5.x 3.4.x 3.3.x
Host Integration Tools for VMware	3.1.2	3.1.x 3.0.x	3.1.x 3.0.x 2.0.x	3.1.x 3.0.x 2.0.x	2.0.x
EqualLogic Storage Replication Adapter for VMware Site Recovery Manager	2.1.x 1.0.6	2.1.x 2.0.x 1.0.x	2.1.x 2.0.x 1.0.x	2.1.x 2.0.x 1.0.x	1.0.x
EqualLogic Multipathing Extension Module for VMware vSphere	1.1.1 1.1.0	1.1.x 1.0.x	1.1.x 1.0.x	1.1.x 1.0.x	1.1.x 1.0.x
Host Integration Tools for Linux	1.1.x	1.1.x 1.0.x	1.1.x 1.0.x	1.1.x 1.0.x	1.0.x
Manual Transfer Utility	1.2.3 1.2.1	1.2.1	1.2.1	1.2.1 1.1.2 (Windows only) 1.0.2 (Linux only)	1.2.1 1.1.2 (Windows only) 1.0.2 (Linux only)

2.3 Optimizing for High Availability and preparing for Array Firmware updates

Business critical data centers must be designed to sustain various types of service interruptions. In an EqualLogic SAN environment, software, servers, NICs, switches, and storage are all interdependent variables that need tuning and configuration with best practices in order to ensure high availability (HA) and non-disruptive operations.

Plan a maintenance window for upgrades if all configuration is not done for proper failover.

Causes of service disruptions:

- Hardware failure – Server, NIC, Cable, Switch, Storage, Controller, Disk, and others
- Software failure – Bugs or Defects, Application Crashes, and others
- Firmware upgrades – Software Patches, Driver Updates, Firmware Updates

A redundant design will survive most of the possible types of service disruptions, but in some cases redundancy alone is not enough. The iSCSI data path is the most crucial point of failure in an EqualLogic SAN environment, and is the focus of this section.

Configuring and tuning your environment with best practice recommendations to design for High Availability and Non-Disruptive Operations can prevent:

- connections to volumes dropping
- application timeout
- group and/or array failure

2.3.1 Verify hosts are designed for high availability

Designing and configuring hosts for high availability helps minimize the chance of disruptions. Recommended settings and configuration options on hosts running in the EqualLogic environment are detailed below:

- Install EqualLogic Host Integration Tools (HIT) on all supported operating systems.
- The host software modules change recommended settings for iSCSI Disk Timeouts.

2.3.2 Microsoft Windows Servers

The Windows Disk I/O timeout parameter is set automatically when EqualLogic HIT for Windows is installed. The TimeOutValue Parameter can be increased to avoid disk I/O timeouts. Verify that the Windows Disk I/O timeout parameter in the Registry Editor is set to no less than 60 seconds.

2.3.3 VMWare ESX Servers

In vSphere 5.x, VMware allocates five seconds for an iSCSI session to login to an iSCSI target. In a normal environment this period of time is sufficient. However, it has been observed that this five second timeout is not long enough to process the large number of iSCSI login requests that occur after a network failure, switch reboot, or controller failover.

It is recommend that the iSCSI Login Timeout value be increased to 60 seconds.

2.3.4 Storage Heartbeat on vSphere 5.0, 4.1, and 4.0

Note: This recommendation for using Storage Heartbeat applies only vSphere 4.1 and 5.0. It is not necessary with vSphere 5.1.

In the VMware virtual networking model, certain types of vmkernel network traffic is sent out on a default vmkernel port for each subnet. If the physical NIC that is being used as the uplink for the default vmkernel port goes down, network traffic that is using the default vmkernel port will fail, including vMotion traffic, SSH access, and ICMP ping replies.

Even though iSCSI traffic is not directly affected by this condition, a side effect of the suppressed ping replies is that the EqualLogic PS Series Group will not be able to accurately determine connectivity during the login process, and therefore a suboptimal placement of iSCSI sessions can occur. This could cause logins to not be completed in a timely manner.

It is recommended to have a high availability vmkernel port on the iSCSI subnet serving as the default vmkernel port for such outgoing traffic.

2.3.5 Linux servers

The HIT Kit for Linux automatically implements recommended changes on the Linux hosts, either at the time of installation or through the built-in "eqltune" utility

2.3.6 Network topology health check

Design your network for redundancy according to recommended best practices:

- For all members (arrays) in a given SAN Group all ports should be connected to the same subnet.
- Use at least two iSCSI SAN ports per host connected, so that at least two of the host ports connect to two different switches in the SAN
- At a minimum Eth0 from Controller0 and Controller1 on each array in the SAN group must be connected to two different switches in the same SAN subnet.
- All switches within the SAN must be interconnected such that there is always a path from any Ethernet port on one array to all other Ethernet ports on all other arrays in the group.

Proper settings configured on switches:

- All switches and host network controllers within the infrastructure must have flow control enabled for optimal performance. For switches, RX flow control is required. For iSCSI hosts, both TX and RX is required.
- Rapid Spanning Tree must be enabled on all switch ports not directly connected to host or array ports.
- Enable PortFast on all switch ports connected to the host or array ports.

See section 7 for complete iSCSI SAN design recommendations.

2.3.7 Environmental health

Verify there are no current health and/or performance problems with the EqualLogic environment.

3 RAID policies

Each array in an EqualLogic array group is configured with a single RAID policy. Arrays (or group members) within the same storage pool that have the same RAID policy will cooperatively work to host volumes by distributing those volumes over multiple arrays. Two things that are defined by the RAID policy are:

- RAID level
- hot-spare configuration

Each array implements a default RAID policy that includes a hot-spare. To configure a RAID policy that does not include a hot-spare, you must manually initialize the array using the Command Line Interface.

3.1 Setting the RAID Policy for a member

PS Series arrays protect data by using RAID technology and spare drives. After you add a member to a PS Series group, you must set the RAID policy for the member. In multi-member groups, you can assign different RAID levels to individual group members.

Once a RAID policy has been specified, the member automatically configures the disk drives according to the designated RAID level and assigns the appropriate number of spare drives. The storage in the member is available after you set the RAID policy, although array performance is not optimal until the system has completed verification of the new RAID configuration.

The Remote Setup Wizard specifies the RAID policy for members when it creates a group and adds the first member to it. See the Installation Guide for your array for information.

If you use the setup utility, you must manually configure the RAID policy as a separate step, using either the Group Manager GUI or the CLI. See the online help or the Dell EqualLogic Group Manager CLI Reference Guide for more information.

Note: Dell recommends against using RAID 5 for any business-critical data, although it may be required for certain applications, depending on performance and data availability requirements. RAID 5 can only be configured using the CLI.

Note: Configuring a RAID policy that does not include a hot-spare will increase the risk of data loss in the event of multiple drive failures.

3.2 Guidelines for choosing a RAID policy

When specifying the RAID policy for a member, the following factors must be considered

3.2.1 RAID level

Each RAID level offers varying combinations of performance, data protection and capacity utilization. Choose your RAID preference carefully based on reliability, capacity, and performance requirements. PS Series arrays support the following RAID types:

- RAID 5 (not recommended for business-critical data)
- RAID 6
- RAID 6 Accelerated
- RAID 10
- RAID 50

See RAID Level Characteristics in Table 8 below for detailed information about the RAID types and their recommended uses.

3.2.2 Drive configuration

The type of drive used in your system factors into choosing an optimal RAID configuration.
The types of drives used in EqualLogic PS Series arrays are listed below.

Class 0 drives are Solid State Drives (SSDs) that offer the highest level of performance. Typically, SSD have smaller capacities than Class 1 or 2 drives.
Class 1 drives are 10,000 and 15,000 RPM SAS hard drives that offer higher performance and greater reliability than Class 2 drives.
Class 2 drives are slower 7200 RPM SATA or Near-Line SAS hard drives that offer greater capacities than Class 1.

If you do not know what types of drives are used in your array, see the Dell EqualLogic PS Series Storage Arrays Release Notes, which lists the types of drives used with each PS Series model. In addition, there is labeling on each drive to indicate its capacity and type.

3.2.3 Spare drive policy

Spare drives increase availability by providing redundancy in the event of a hard disk drive failure or a solid-state disk drive failure. The number of spare drives depends on the array model and the number of installed drives. For RAID 6 Accelerated, only one hard disk drive is configured as a spare. Spare HDD drives can be used as spares for either HDD or SSD drives.
Dell does not recommend using any configuration that does not utilize spare drives, although these configurations can be implemented using the CLI. For more information about using the CLI, refer to the Dell EqualLogic Group Manager CLI Reference Guide.

3.3 RAID level characteristics

Table 8 summarizes best practices for RAID policy selection. When selecting a RAID policy, you should consider the level of data protection you require, as well as I/O performance, capacity needs, and the types of drives contained in your PS Series array.

Table 8 RAID level characteristics

Raid policy	Recommended usage scenarios	Recommended drive configurations
RAID 10	Applications and workloads requiring the highest levels of I/O performance for random writes.	Systems containing 10K and 15K RPM drives.
RAID 6	Situations in which I/O performance for random writes is not a key factor. Applications requiring the highest levels of data protection and reliability.	Systems containing 24 or more drives. Systems containing 7200 RPM SATA or Nearline SAS (NL-SAS) drives.
RAID 6 Accelerated	Optimizes the use of solid-state drives for critical data.	Supported only on arrays that contain both solid state (SSD) and hard disk (HDD) drives. For these systems, it is the only available RAID policy. It is not a user-selectable option.
RAID 50	Applications requiring an optimal combination of performance, availability, and capacity.	Systems containing fewer than 24 drives. Systems containing 10K and 15K RPM drives.
RAID 5	Dell recommends against using RAID 5 for any business-critical data, although it may be required for certain applications, depending on performance and data availability requirements. RAID 5 carries higher risks of encountering an uncorrectable drive error during a rebuild, and therefore does not offer optimal data protection. RAID 5 can only be configured using the CLI.	N/A

3.3.1 Performance and availability characteristics of the supported RAID levels

Table 9 compares the performance and availability characteristics of the supported RAID levels. The first column lists workload requirements, with the other columns respectively listing each RAID level's characteristics with respect to each requirement.

Table 9 RAID Level Characteristic Comparison Workload Requirement

Workload Requirement	RAID 5	RAID 6	RAID 10	RAID 50
Capacity	Excellent	Good	Average	Good
Availability	Poor	Excellent	Good	Fair
Sequential reads	Excellent	Excellent	Excellent	Excellent
Sequential writes	Good	Good	Good	Good
Random reads	Excellent	Excellent	Excellent	Excellent
Random writes	Good	Fair	Excellent	Good
Performance impact of drive failure or RAID reconstruction	Longest	Longest	Shortest	Medium

3.4 Supported RAID policy conversions

While a member remains online, you can convert it from one RAID policy to another only if the new RAID policy provides the same or more space than the current policy.

The table below shows the supported RAID policy conversions. The first column lists RAID configurations, and the second column lists supported RAID policy conversions for them.

Table 10 Supported RAID policy conversions

Current RAID policy	Supported conversion
RAID 10	All
RAID 50	RAID 5, RAID 6
RAID 5	None
RAID 6	None
RAID 6 Accelerated	None

If a RAID policy conversion is not supported, you can remove the member from the group and then add it to the group again. You can then set the RAID policy.

Before changing a member's RAID policy, note the following:

- Single member groups: If you change the RAID policy in a single member group, all data on the member will be lost.
- Multi-member groups: Before changing a group member's RAID policy, you must first vacate its data to another member in the group, or all data on the member will be lost. When you vacate a member, there must be sufficient available free space on another member to store the evacuated data.

To convert from a no-spares RAID policy to a policy that uses spare drives, you must use the CLI. Refer to the Dell EqualLogic Group Manager CLI Reference Guide. You can also use the CLI to convert to a RAID policy that does not use spare drives, but Dell recommends against doing this.

4 Capacity planning

4.1 RAID 6 drive layouts and total reported usable storage

RAID6 (striped set with dual distributed parity) combines N disks in an arrangement where each stripe consists of N-2 disks capacity for data blocks and two disks capacity for parity blocks. Each parity block generates parity using a different view of the data blocks depending on the RAID 6 implementation. RAID 6 can tolerate up to two drive failures per RAID stripe set at the same time without data loss. RAID 6 is not recommended for workloads consisting mainly of random writes. Using a RAID 6 policy, Table 11 shows the drive layouts that are enforced based on the number of drives in each array/hot spare configuration, and the total usable storage available for each model.

Table 11 RAID 6 drive layouts and total storage available with hot spares (in GB)

Disk drives	Hot spare	No hot spare
6	5 Data/Parity + 1 Hot-spare	6 Data/Parity
7	6 Data/Parity + 1 Hot-spare	7 Data/Parity
8	7 Data/Parity + 1 Hot-spare	8 Data/Parity
12[f]	11 Data/Parity + 1 Hot-spare	12 Data/Parity
14	13 Data/ Parity + 1 Hot-spare	14 Data/Parity
16	15 Data/ Parity + 1 Hot-spare	16 Data/Parity
8 + 8[c]	15 Data/Parity(d) + 1 Hot-spare	16 Data/Parity
24[a][f]	23 Data/Parity + 1 Hot-spare	24 Data/Parity
7 + 17[e]	23 Data/Parity + 1 Hot-spare	24 Data/Parity
48[a]	47 Data/Parity + 1 Hot-spare	N/A[b]

(a) 24 and 48 drive arrays implement multiple RAID 6 sets within a single chassis.
(b) 48 drive arrays cannot implement a no hot-spare RAID policy.
(c) PS60x0 XVS with 8x 100GB SSD and 8x 450GB SAS.
(d) One 450GB hot spare, 2x 450 GB parity and 2x 100GB SSD parity. In the event of a drive failure (SSD or SAS), the RAID set is reconstructed using the 450GB SAS hot spare.
(e) One 600GB hot spare, 2x 600 GB parity and 2x 400 GB SSD parity. In the event of a drive failure (SSD or SAS), the RAID set is reconstructed using the 600 GB SAS hot spare.
(f) PS4100/PS6100 models only

Total reported usable storage when using hot spares: All models prior to PS4100/PS6100

Drive Qty / Size	50[a]	100[a]	74	146	250	300	400	450	500	600	750	1000	2000	3000
7[b]	-	-	-	-	888	-	-	-	-	-	-	-	-	-
8[c]	-	-	-	-	1111	-	-	-	-	-	-	-	-	-
14[d]	-	-	660	-	2224	-	3566	-	4460	-	6695	-	-	-
16	577	1155	-	1696	2895	3475	4638	5220	5800	6964	8708	11613	23633	35717
8 + 8[e]	-	-	-	-	-	-	-	2539	-	-	-	-	-	-
48	-	-	-	-	-	-	-	-	18062	21677	-	36136	63637	107151

Total reported usable storage when using hot spares: PS41x0/PS61x0

Drive Qty / Size	146	200	300	400	500	600	900	1000	2000	3000
6	-	-	-	-	1331	1597	-	2672	5457	8253
12	1177	1658	2406	3328	4014	4812	7229	8038	16363	24719
14	1434	-	2938	-	4906	5890, 4286[h]	8841	9824	-	-
24	2478	3512[f]	5084	7034[f]	8478	10178	15278	16947	34550	52202
7 + 17	-	-	-	-	-	9318[g]	-	-	-	-
48	-	-	-	-	-	-	-	-	63637	-

(a) Solid State Disk
(b) Legacy PS50E
(c) PS4000E and PS6000E only
(d) Legacy PS70E, PS100E, PS100X, PS200E, PS300E and PS400E
(e) PS60x0 XVS with 8x100GB SSD and 8x450GB SAS
(f) PS6100S (SSD) only
(g) PS6100XS only
(h) M4110XS only

4.2 RAID 10 drive layouts and total reported usable storage

Using a RAID 10 policy, Table 12 shows the drive layouts that are enforced based on the number of drives in each array/hot spare configuration, and the total usable storage available for each model.

RAID 10 (mirrored sets in a striped set) combines two high performance RAID types: RAID 0 and RAID 1. A RAID 10 is created by first building a series of two disk RAID 1 mirrored sets, and then distributing data over those mirrors. RAID 10 implementations can tolerate one drive failure per mirror pair.

Table 12 RAID 10 drive layouts and total reported usable storage available with hot spares (in GB)

Disk drives	Hot spare	No hot spare
6	4 Data (2 mirrors) + 2 Hot-spares	6 Data (3 mirrors) + 0 Hot-spare
7	6 Data (3 mirrors) + 1 Hot-spare	N/A[a]
8	6 Data (3 mirrors) + 2 Hot-spares	8 data (4 mirrors)
12[c]	10 Data (5 mirrors) + 2 Hot-spares	12 data (7 mirrors)
14	12 Data (6 mirrors) + 2 Hot-spares	14 data (7 mirrors)
16	14 Data (7 mirrors) + 2 Hot-spares	16 data (8 mirrors)
24[c]	22 Data (11 mirrors) + 2 Hot-spares	24 data (12 mirrors)
48	46 Data (23 mirrors) + 2 Hot-spares	N/A[b]

(a) Not supported. An even number of drives is required by RAID 10
(b) 48 drive arrays cannot implement a no hot-spare policy.
(c) PS4100/PS6100 models only

Total reported usable storage when using hot spares: All models prior to PS4100/PS6100

Disk drive qty / Size	50[a]	100[a]	74	146	250	300	400	450	500	600	750	1000	2000	3000
7[b]	-	-	-	-	666	-	-	-	-	-	-	-	-	-
8[c]	-	-	-	-	666	-	-	-	-	-	-	-	-	-
14[d]	-	-	395	-	1333	-	2135	-	2671	-	4013	-	-	-
16	309	621	-	913	1556	1867	2492	2805	3117	3744	4683	6247	12749	19230
48	-	-	-	-	-	-	-	-	10647	12779	-	21306	40840	63191

Total reported usable storage when using hot spares: PS41x0/PS61x0

Disk drive qty / Size	146	200[e]	300	400[e]	500	600	900	1000	2000	3000
6	-	-	-	-	887	1064	-	1781	3635	5498
12	650	920	1331	1843	2222	2672	4014	4454	9103	13762
14	781	-	1599	-	2670	3207	4816	5353	-	-
24	1433	2027	2938	4065	4904	5888	8841	9820	19998	30221

(a) Solid State Disk
(b) Legacy PS50E
(c) PS4000E and PS6000E only
(d) Legacy PS70E,PS100E, PS100X, PS200E and PS300E
(e) PS6100S (SSD) only

4.3 RAID 50 drive layouts and total reported usable storage

Table 13 shows the drive layouts that are enforced when using a RAID 50 policy based on the number of drives in each array/hot spare configuration and the total usable storage available for each model.

RAID 50 (RAID 5 sets in a striped set) is created by first creating two or more RAID 5 sets and then striping data over those RAID5 sets. RAID 50 implementations can tolerate a single drive failure per RAID5 set.

Table 13 RAID 50 drive layouts and total reported usable storage available with hot spares (in GB)

Disk drives	Hot spare	No hot spare
6	5 Data/Parity + 1 Hot-spare	N/A[a]
7	6 Data/Parity + 1 Hot-spare	N/A[a]
8	6 Data/Parity + 2 Hot-spares	8 data/parity
12[d]	10 Data/Parity + 2 Hot-spares	12 data/parity
14	12 Data/Parity + 2 Hot-spares	14 data/parity
16	14 Data/Parity + 2 Hot-spares	16 data/parity
24[d]	22 Data/Parity + 2 Hot-spares	24 data/parity
48[b]	46 Data/Parity + 2 Hot-spares	N/A[c]

(a) RAID 50 requires an even number of disk drives. A 7 drive configuration without hot-spare would result in odd number of disk drives.
(b) 48 drive arrays implement stripes across multiple RAID 5 sets within a single chassis.
(c) 48 drive arrays cannot implement a no hot-spare policy.
(d) PS4100/PS6100 models only

Total reported usable storage when using hot spares: All models prior to PS4100/PS6100

Disk Drive Qty / Size	50[a]	100[a]	74	146	250	300	400	450	500	600	750	1000	2000	3000
7[b]	-	-	-	-	888	-	-	-	-	-	-	-	-	-
8[c]	-	-	-	-	888	-	-	-	-	-	-	-	-	-
14[d]	-	-	660	-	2224	-	3566	-	4460	-	6695	-	-	-
16	532	1066	-	1566	2671	3207	4280	4817	5353	6427	8037	10719	21819	32972
48	-	-	-	-	-	-	-	-	18062	21677	-	36136	69260	107151

Total reported usable storage when using hot spares: PS41x0/PS61x0

Disk Drive Qty / Size	146	200[e]	300	400[e]	500	600	900	1000	2000	3000
6	-	-	-	-	1781	2129	-	3563	7280	11008
12	1044	-	2129	2949	3563	4280	6430	7137	14571	21975
14	1304	-	2670	-	4459	5353	8036	8930	-	-
24	2355	3328	4815	6666	8038	9646	14474	16087	32727	49455

(a) Solid State Disk
(b) Legacy PS50E
(c) PS4000E and PS6000E only
(d) Legacy PS70E,PS100E, PS100X, PS200E and PS300E
(e) PS6100S (SSD) only

4.4 RAID 5 drive layouts and total reported usable storage

RAID 5 (striped disks with distributed parity) will combine N disks in an arrangement where each stripe consists of N–1 disks that contain data blocks plus 1 disk that contains a parity block. For each stripe, the parity block will be placed on a different disk ensuring that the parity blocks are not located on a single disk in the RAID set. RAID 5 implementations can tolerate a single drive failure without data loss.

Table 14 shows the drive layouts that are enforced when using a RAID 5 policy based on the number of drives in each array/hot spare configuration, and the total usable storage available for each model.

Table 14 RAID 5 drive layouts and reported usable storage available with hot spares (in GB)

Disk drives	Hot spare	No hot spare
6	5 Data/Parity + 1 Hot Spare	RAID Policy for RAID 5 without a hot spare is not supported
7	6 Data/Parity + 1 Hot-spare	
8	7 Data/Parity + 1 Hot-spare	
12[a]	11 Data/Parity + 1 Hot-spare	
14	13 Data/Parity + 1 Hot-spare	
16	15 Data/Parity + 1 Hot-spare	
24[a][b]	23 Data/Parity + 1 Hot-spare	
48[b]	46 Data/Parity + 2 Hot-spares	

(a) PS4100/PS6100 models only
(b) 24 and 48 drive arrays implement multiple RAID 5 sets within a chassis

Total reported usable storage when using hot spares: All models prior to PS4100/PS6100

Drive qty / Size	50[a]	100[a]	74	146	250	300	400	450	500	600	750	1000	2000	3000
7[b]	-	-	-	-	1111	-	-	-	-	-	-	-	-	-
8[c]	-	-	-	-	1333	-	-	-	-	-	-	-	-	-
14[d]	-	-	792	-	2671	-	4281	-	5354	-	-	-	-	-
16	621	1244	-	1827	3118	3744	4995	5622	6247	7500	9378	12508	25456	38461
48	-	-	-	-	-	-	-	-	19452	23345	-	38916	74580	115394

Total reported usable storage when using hot spares: PS41x0/PS61x0

Drive qty / Size	146	200[e]	300	400[e]	500	600	900	1000	2000	3000
6	-	-	-	-	1781	2129	-	3563	7280	11008
12	1300	1843	2672	3696	4454	5355	8038	8929	18186	27473
14	1565	-	3206	-	5353	6426	9645	10719	-	-
24	2744	3880	5620	7772	9379	11253	16855	18728	38184	57698

(a) Solid State Disk
(b) Legacy PS50E
(c) PS4000E and PS6000E only
(d) Legacy PS70E,PS100E, PS100X, PS200E, and PS300E
(e) PS6100S (SSD) only

4.5 Array RAID configurations and associated RAID sets

The tables show a logical drive layout when an array is initialized for the first time. The actual physical layout of drives can change and evolve due to maintenance and administrative actions. Spare drives can move as they are utilized to replace failed drives and newly added drives become the spares. It is not possible to determine which physical drives are associated with each RAID set. This information is dynamic and maintained by the EqualLogic firmware.

Table 15 shows the RAID set relationship for each RAID type in a 24-drive configuration.

Table 15 EqualLogic PS Series array RAID types and RAID set relationships

RAID policy	Spare disks	RAID set relationship	Best practice
RAID 6	1 Spare Disk	(10+2) (9+2)	Yes
RAID 10	2 Spare Disks	(6+6) (5+5)	Yes
RAID 50	2 Spare Disks	(5+1, 5+1) (4+1, 4+1)	For selected configurations
RAID 5	1 Spare Disk	(12+1) (9+1)	Not for business critical data

Table 16 shows the RAID set relationship for each RAID type in a 48-drive configuration.

Table 16 EqualLogic PS Series array RAID types and RAID set relationships

Raid policy	Spare disks	Raid set relationship	Best practice
RAID 6	2 Spare	(12+2, 12+2, 12+2) (2+2)	Yes
RAID 10	2 Spare	(7+7, 7+7, 7+7) (2+2)	Yes
RAID 50	2 Spare	(6+1 ,6+,1, 6+1, 6+1, 6+1, 6+1) (3+1)	For selected
RAID 5	2 Spare	(12+1, 12+1, 12+1) (6+1)	Not for business critical

Table 17 shows the RAID set relationship for each RAID type in a 16-drive configuration.

Table 17 EqualLogic PS Series array RAID types and RAID set relationships

Raid policy	Spare disks	Raid set relationship	Best practice
RAID 6	1 Spare Disk	(13+2)	Yes
RAID 10	2 Spare Disks	(7+7)	Yes
RAID 50	2 Spare Disks	(6+1, 6+1)	For selected configurations
RAID 5	1 Spare Disk	(14+1)	Not for business critical data

5 PS Series array concepts

5.1 Groups and pools

A PS Series SAN Group is a Storage Area Network (SAN) comprised of one or more PS Series arrays connected to an IP network. Each array in a group is called a group *member*. Each member is assigned to a storage pool. There can be up to 4 pools within the group.

A group can consist of up to 16 arrays of any family or model as long as all arrays in the group are running firmware with the same major and minor release number. For example, it is supported to have different arrays in the same group running different maintenance release levels, as long as their major and minor revision levels match. An example of this would be running version 4.0.1 and 4.0.3 on different arrays in the same group. An example of an unsupported configuration would be different arrays in the same group running 4.0.1 and 4.1.0 (the minor versions differ). The only exception to this rule is for short term time periods when array firmware versions may be out of sync while upgrading the firmware on each array within a group. Features available for use within the group are determined by the lowest version of FW running within the group.

> **Note:** We recommended that all arrays run the same version of PS Series firmware at all times, except during a firmware upgrade process. (See Section 2.1 for more information.)

5.1.1 Pools

A pool is a container that each member is assigned after being added to the group. A pool can have between one to eight members. There is always at least one pool in any group and it is called the "default pool" unless the name is changed. Regardless of the name of this pool, it is always considered the default storage pool. All newly added members are automatically assigned to the default pool. The default pool cannot be deleted.

Pools can be used as part of an overall tiered storage strategy for the SAN. Tiered storage is storage that is differentiated and optimized for the type of applications and data being hosted. Instead of optimizing all of the storage in the group for maximum storage utilization, a tiered storage system allows for the administrator to optimize different arrays for a variety of requirements such as application performance requirements or cost efficiency.

Pools are the containers that provide scope for all automated storage optimization features of the group. Pools with multiple arrays can implement different RAID policies within the pool. The EqualLogic automated optimization features will automatically move a volume from an array or set of arrays with one RAID policy to another array or set of arrays with a different RAID policy if it is determined that the application needs more (or less) performance.

The following rules apply to pools:

- Each member can be assigned to only one pool.
- Each pool can host up to eight members.
- Each group will have at least one pool – the default pool.
- Each group can have up to four pools.

- If all members in the pool are running PS Series firmware v5.0 or later then you can mix PS5500E, PS6500E/X and PS6510E/X models together with other array models in the same pool.
- If you are running PS Series firmware version prior to v5.0 then PS5500E, PS6500E/X and PS6510E/X models must reside in a separate pool from other array types.

Figure 3 shows a PS Series group with the maximum of four pools. Note the use of Pool 3 for containing PS5500/PS6500 series arrays only. Also note that Pool 3 contains arrays that implement both 1GbE and 10GbE controllers.

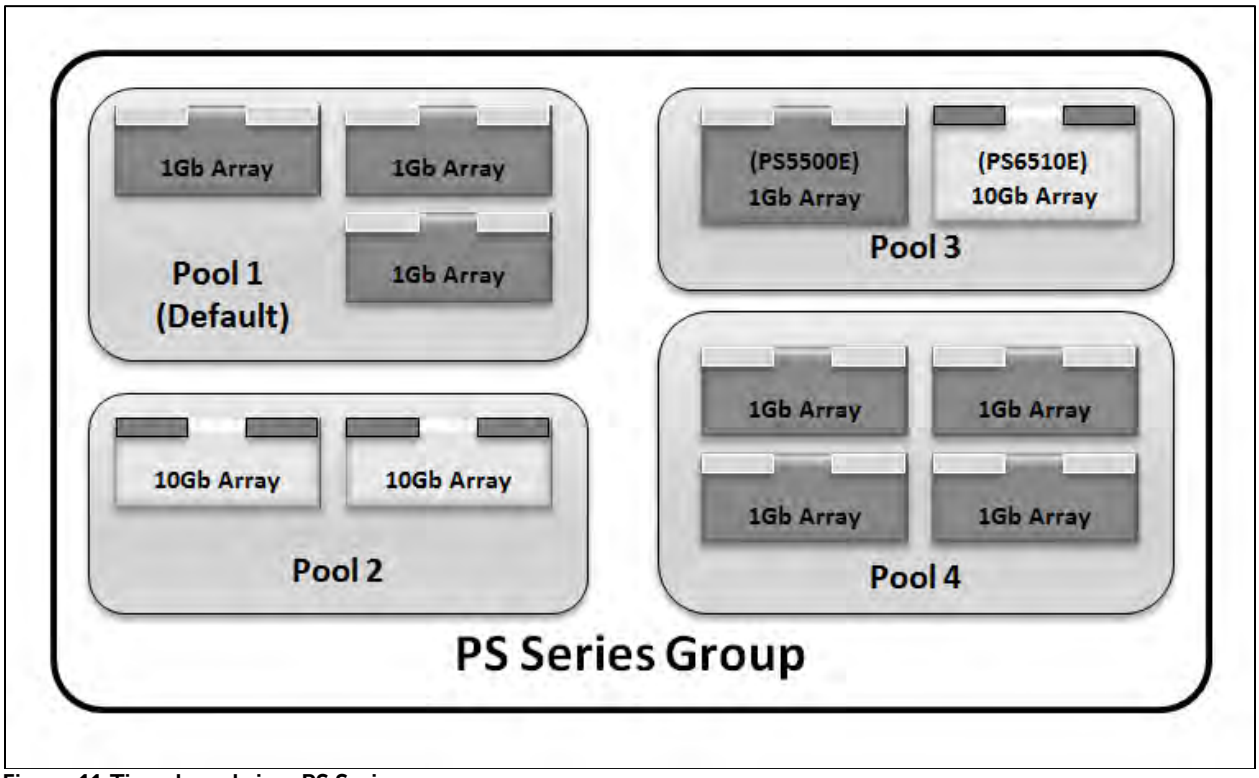


Figure 11 Tiered pools in a PS Series group

The following best practices should be considered for storage pools unless the automatic performance load balancer (APLB) is disabled:

- Mixing arrays of different drive speeds and types is fully supported and is the primary reason APLB was created.
- Do not mix arrays with different controller speeds (1GbE, 10GbE) within a single pool unless they are each running unique RAID policies.
- To override the automated performance decisions for a specific volume, indicate a "preferred" RAID type for that volume. If that RAID type exists within the pool, the volume will reside on those arrays that match the preferred RAID policy for the volume.

If the APLB is disabled, the following recommendations apply:

- Do not mix arrays with different drive speeds within a single pool unless they are running a unique RAID policy

- Do not mix arrays with different drive technologies (SATA, SAS, SSD) within a single pool unless they are running a unique RAID policy.
- Do not mix arrays with different controller speeds (1GbE, 10GbE) within a single pool unless they are each running unique RAID policies.

5.2 Volumes

Volumes provide the storage allocation structure within an EqualLogic SAN. Volumes are seen on the network as iSCSI targets by hosts and are presented to the user as disk drives once the iSCSI initiator has been configured and authenticated to the volume. Only computers with an iSCSI initiator and the correct access credentials can access a volume. Disk space for volumes is allocated from the target storage pool's free space.

Volumes are identified within the Group Manager with a unique volume name. The volume name is created by the administrator during volume creation and can be configured with several additional characteristics and capabilities. The following sections provide additional details.

5.2.1 Volume attributes

Volumes are created using the "create volume" function in Group Manager or through the Command Line Interface (CLI). Each volume must have a unique name that is used by Group Manager to identify and manage the volume. Volume names may be changed via Group Manager at any time. Volume names must meet the following requirements:

- 1 to 64 alpha-numeric characters
- A-Z, a-z, 0-9, ".", "-", ":" are legal characters

Volumes are assigned a unique iSCSI Qualified Name (iqn) that consists of the following parts:

- "**iqn**" followed by a "**.**" (period)
- Year and Month of first full month that the naming authority was registered with standards body. EqualLogic's value is "**2001-5**" followed by a "**.**"
- The storage provider's domain name in reverse order. For example: "**com.equallogic**"
- A colon ("**:**")
- Vendor specified information to identify an iSCSI target or initiator as unique within the vendor domain. For EqualLogic iSCSI targets, this part consists of a unique set of numbers and the user assigned volume name (within Group Manager).

The following is an example of an iSCSI target name for a volume named db3:

```
iqn.2001-05.com.equallogic:7-8b0900-6d0000000-001ebbc5d80sf0k0-db3
```

Volumes are visible to iSCSI initiators through one or more of three mechanisms during volume creation:

- iSCSI initiator IP address
- A host's iSCSI iqn well-known name
- Mutual CHAP authenticated user identity

Note: IQN names are assigned to volumes automatically when they are created. They cannot be changed for the life of the volume. If a volume name is changed, the IQN name associated with the volume will remain unchanged.

5.2.2 Volume features

Each volume supports a set of features and capabilities that include the following:

- Ability to define a volume as thin-provisioned
- Support for snapshots
- Support for replication
- Support for creating clones
- Support for multiple host shared access

Note: These block level volume features are not applicable to the component volumes created by the NAS Service for the FS Series NAS Reserve.

5.3 Snapshots and clones

Snapshots are point in time copies of volumes. Snapshots have some features and properties similar to volumes as well as some unique capabilities. Like volumes, snapshots can be assigned an iqn and presented as volumes to a host. This allows hosts to mount a snapshot, potentially change the data in the snapshot, or convert it to a real volume (clone) that has dedicated space allocated to it from the free storage pool.

Snapshots require that space be reserved during volume creation (or after volume is created) to store the snapshot(s) created from that volume. All snapshot reserve space for a specific volume always resides in the same storage pool as the volume. By default, Group Manager allocates snapshot reserve space equal to 100% of the host volume's allocated space. This ensures that a 100% data change in the volume can be protected by a single snapshot. This value can be set to a lower value based on the application's data change rate, snapshot/backup plan, or role that the snapshot will be used. The following items should be considered when determine the size of a volume's snapshot reserve:

- Data change rate for the application(s) who is using the volume
- Defragmentation policy of the OS or application using the volume
- The role or purpose for creating the snapshot(s).

Snapshots have the following characteristics:

- They are identified using names that consist of the parent volume name plus a date/time stamp indicating when the snapshot was created as a default name.
- If name of parent volume changes, existing snapshots iqn names do NOT change accordingly, but retain their original iqn name.
- Deleting a snapshot's parent volume deletes all associated snapshots for that volume.
- Snapshots of volumes with a high data change rate will require a larger snapshot reserve space.
- Snapshots have access control lists that are inherited from the parent volume by default.

- Snapshot reserve space for any volume can be decreased at any time. The minimum size allowed is based on the current space used by existing snapshots using the snapshot reserve.
- Snapshot reserved space for any volume can be increased at any time assuming there is available free space in the storage pool hosting the volume.

5.3.1 Clones

Cloning creates a new volume by copying an existing volume. The new volume has the same reported size, contents and thin-provision settings as the original volume. You can clone a regular volume, a specific replica of a volume, or a specific snapshot of a volume.

- Volume clones use non-shared storage (unlike snapshots)
- Volume clones reduce the amount of storage pool free space
- All regular volume operations can be performed on the clone once the cloning operation is complete.
- The cloning operation supports instant accessibility (while the clone operation is occurring)

Table 18 shows clone sources and properties.

Table 18 Cloning modes and properties

Clone source	Common properties
Clone from Volume	• Creates a new volume • Creates a new volume name • Creates a new iSCSI target IQN • Has same reported size as source volume • Has same contents as source volume • Cloned volume is immediately available • Cloning consumes free pool space equivalent to 100% of the volume reserve setting for the source volume[(a)]
Clone from Snapshot	
Clone from Inbound Replica[(a)] (secondary group)	

(a) Replica clones are created in the secondary group pool and are immediately available at the secondary group IP address.

5.4 Thin provisioning

Thin provisioned volumes provide administrators with an additional option for managing data growth within the SAN. Thin provisioning allows volumes to present a logical capacity to the host(s) connecting to the volume that is different (larger) than the actual physical storage resource allocation used by the volume at any given time.

Note: Volumes automatically created and managed by the NAS Service inside the NAS Reserve pool used by an FS Series appliance are fully allocated at creation and cannot be thin provisioned. Once those volumes are created (when the NAS Reserve is created they cannot be shrunk.

A volume can be provisioned as a thin provisioned volume at creation or after creation. The following rules apply to thin provisioned volumes:

- A minimum physical allocation of 10% of the logical allocation is required
- If a volume is converted to a thin provisioned volume, physical allocation cannot be less than the amount of physical space already used within the volume
- Any pool free space allocated to a thin provisioned volume is not returned to the free pool if the host's file system usage of that volume is reduced (due to file system defragmentation, data removal, etc.)
- If a thin provisioned volume's allocated space exceeds the "maximum in-use space" setting, the volume will go into an offline state.
- Converting a volume to a thin provisioned volume may not reduce the physical allocation of space the volume is using. Actual physical space recovered will depend on previous writing patterns for the OS file system hosted on the volume.
- All initial settings for "minimum volume reserve", "In-use volume reserve warning level" and "maximum in-use space" are default recommendations and can be changed by the administrator within the constraints defined above.

The following best practices should be considered when using thin provisioned volumes:

- Use Pool Free Space, not Group Free Space when making all determinations of thin provisioned volume physical capacity allocation.
- Create regular volumes before creating thin provisioned volumes. This provides the administrator with a better view of the remaining available free space in the pool.
- Set each thin provisioned volume's "In use Volume Reserve Warning Level" to a level that allows the administrator to ensure that additional physical storage can be deployed before the volume uses the remainder of current pool free space.
- Ensure that the Sum of all "In use Volume Reserve Warning Levels" for all thin provisioned volumes does not exceed the current pool free space capacity minus 10%.
- The storage administrator and application owner should reach an agreement on storage use and procedures in determining the best automatic growth policy, reducing the frequency of monitoring and intervention.
- Be aware of the data usage patterns of the application. Some applications perform automated disk optimization that may cause a thin provisioned volume to use more physical storage than would be needed for normal operations. For these types of applications, thin provisioned volumes may not be indicated.
- Use "quick format" options when formatting OS file systems that are hosted by thin provisioned volumes.
- Thin provisioned volumes should not be used for hosting OS boot partitions or OS page file cache.

5.4.1 Template volumes and thin clones

Template volumes and thin clones are new PS Series features introduced with firmware version 5.0. Thin clones are created from template volumes. With a few exceptions, all normal volume operations apply to template volumes and thin clones. For details on using template volumes and thin clones, see the *Advanced volume operations → About template volumes and thin clones* section of the *PS Series Group Manager* documentation.

6 Array firmware features

6.1 Replication

Replication is a powerful feature that can help you manage and implement a disaster recovery strategy for your business applications and data. By replicating business-critical volumes, you ensure that your business operations can resume quickly on a partner group in the event of a disaster on the primary group. You also have the ability to restore the configuration to its original state if the problem on the original group can be corrected. The replication process does not require placing volumes offline. Thus you can take advantage of the benefits of replication without causing interruption in application or data availability.

In EqualLogic PS Series storage, replication is the process of copying volume data from one PS Series group to another PS Series group. Physical distance between the replication groups is not a concern as long as they are connected by a reliable TCP/IP network path.

Table 19 provides definitions of EqualLogic PS Series terminology related to replication processing.

Table 19 Replication terminology

Term	Description
Replica	A time synchronized copy of an EqualLogic volume stored in a Secondary Group.
Replica Set	A collection of all time synchronized replicas for a specific source volume.
Primary Group	A group containing the source volume(s) to be copied or replicated.
Source Group	Same as “Primary Group”.
Secondary Group	A group containing the replica or copy of the source volume(s).
Destination Group	Same as “Secondary Group”.
Delegated Space	The amount of space on the secondary group that is “delegated” to a replication partner, to be reserved for retaining replicas.
Replica Reserve	The space allocated from delegated space in the secondary group to store the volume replica set for a specific volume.
Local Reserve	The amount of space reserved on the local or Primary Group for holding temporary snapshots and failback snapshots of the source volume.

6.1.1 Replication limits

You need to consider the following limits when designing a volume replication topology:

- A primary group can replicate to multiple partner (secondary) replica groups.

- PS Series groups can have up to 16 replication partners and can support a maximum of 10,000 total snapshots and replicas from all of its replication partners.
- A group can have volumes replicating with multiple partners, but an individual volume can have only one replication partner.
- A maximum of 256 volumes per group can be configured for active replication.
- All volumes that are part of a volume collection can only replicate with a single partner.
- A volume can have up to a maximum of 512 replicas stored on a partner group.
    - For a PS4000 only group, you cannot exceed two replication partners and 32 volumes configured for replication

6.1.2 Replication paths

The example replication paths shown in Figure 12 are described below.

- **Basic partnership:** one partner group hosts the primary copy of the volume and a second partner group hosts the replica copy of the same volume. We also show the reverse direction of the path for the Fast Failback replication feature, if it is enabled for the replica set.
- **Single to multiple group partnership:** a single group can support replication of multiple primary volumes to different secondary replication partner groups.
- **Reciprocal partnership:** you can create bi-directional replication paths between partner groups
- **Multiple to single group partnership:** a common scenario in which multiple primary groups replicate their volumes to a single secondary group partner in a different location.

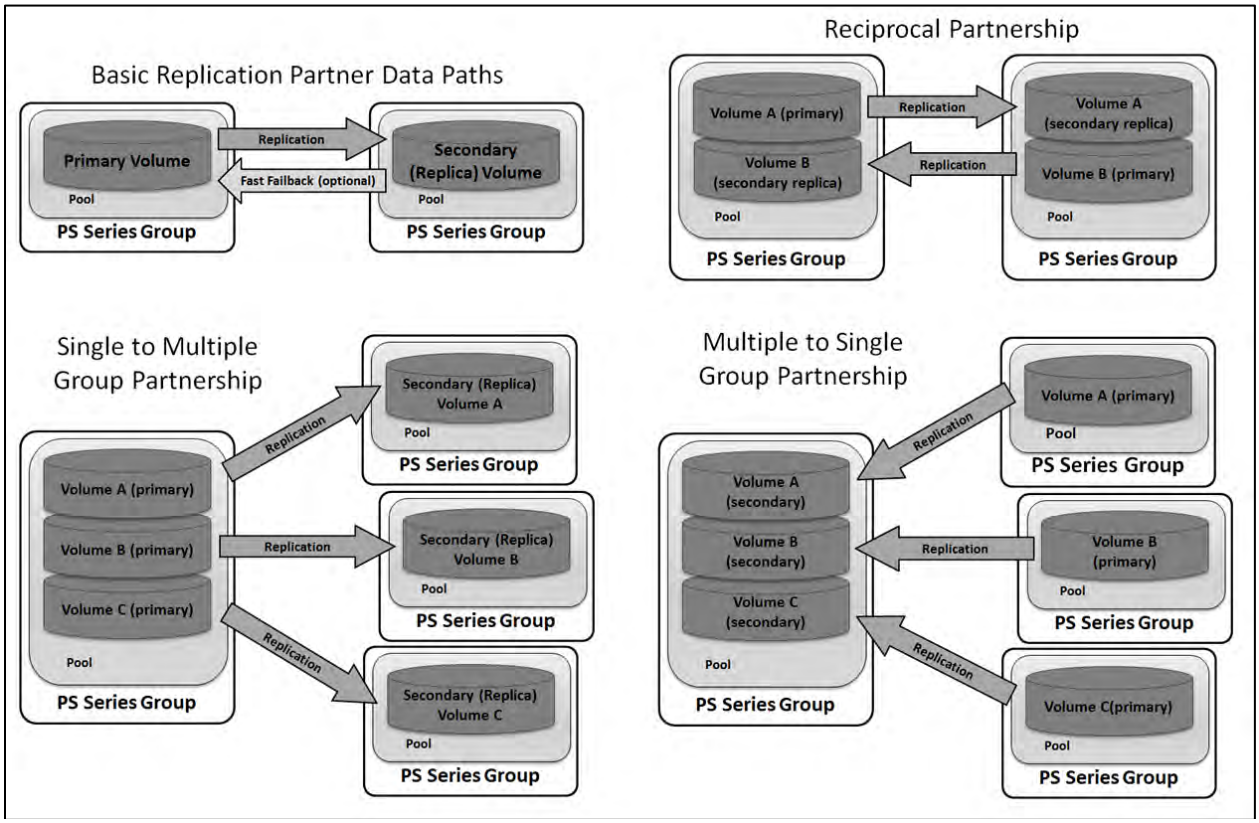


Figure 12 Replication partnership paths

6.1.3 Replication process

When a replica is created the first replication process completes transfer of all volume data. For subsequent replicas, only the data that changed between the start time of the previous replication cycle and the start time of the new replication cycle is transferred to the secondary group. Dedicated volume snapshots are created and deleted in the background as necessary to facilitate the replication process.

A volume replica set is defined as follows:

Volume Replica Set = A full copy of the primary volume, with data synchronized to the beginning of the most current completed replication + A time sequenced set of replicas, where each replica corresponds to the state of the volume at the beginning of a prior replication.

The number of prior replicas in the replica set that can stored on the secondary group is limited by the size of the Replica Reserve allocated for that volume and the amount of data that changes.

Replication processing occurs in a series of phases. The flowchart in Figure 13 shows the process phases, focusing on how the process tracks and copies changes that occur between each replica cycle.

Referring to the phases shown in Figure 13:

Replication Setup (one-time):
Configure replication partnership and volume replication settings.

Replication Processing (repeating):
The primary group checks for availability of sufficient delegated and replica reserve space on the secondary group at the beginning of each replication processing phase. If adequate space is not available then the process will pause and generate an event message. Replication will continue once sufficient space is made available. This part of the process is not shown in the chart.

Primary to secondary volume data replication is completed. The process steps vary based on replication status (first or subsequent) and fast failback mode (enabled or disabled). During this process Local Reserve is consumed by a hidden snapshot (and the fast failback snapshot if enabled). Volume data changes that occur during the replication processing phase are stored by the hidden snapshot in Local Reserve. Replica Reserve allocated to the volume within delegated space on the secondary group receives all volume data changes. Replica Reserve is consumed by the most recent complete volume replica plus all prior replicas stored in the replica set.

Between Replication Events (repeating):
Once first replication has occurred the system continues to keep track of volume data changes that occur so that subsequent replication processes can copy those changes to the replica set. This tracking process does not consume additional space.

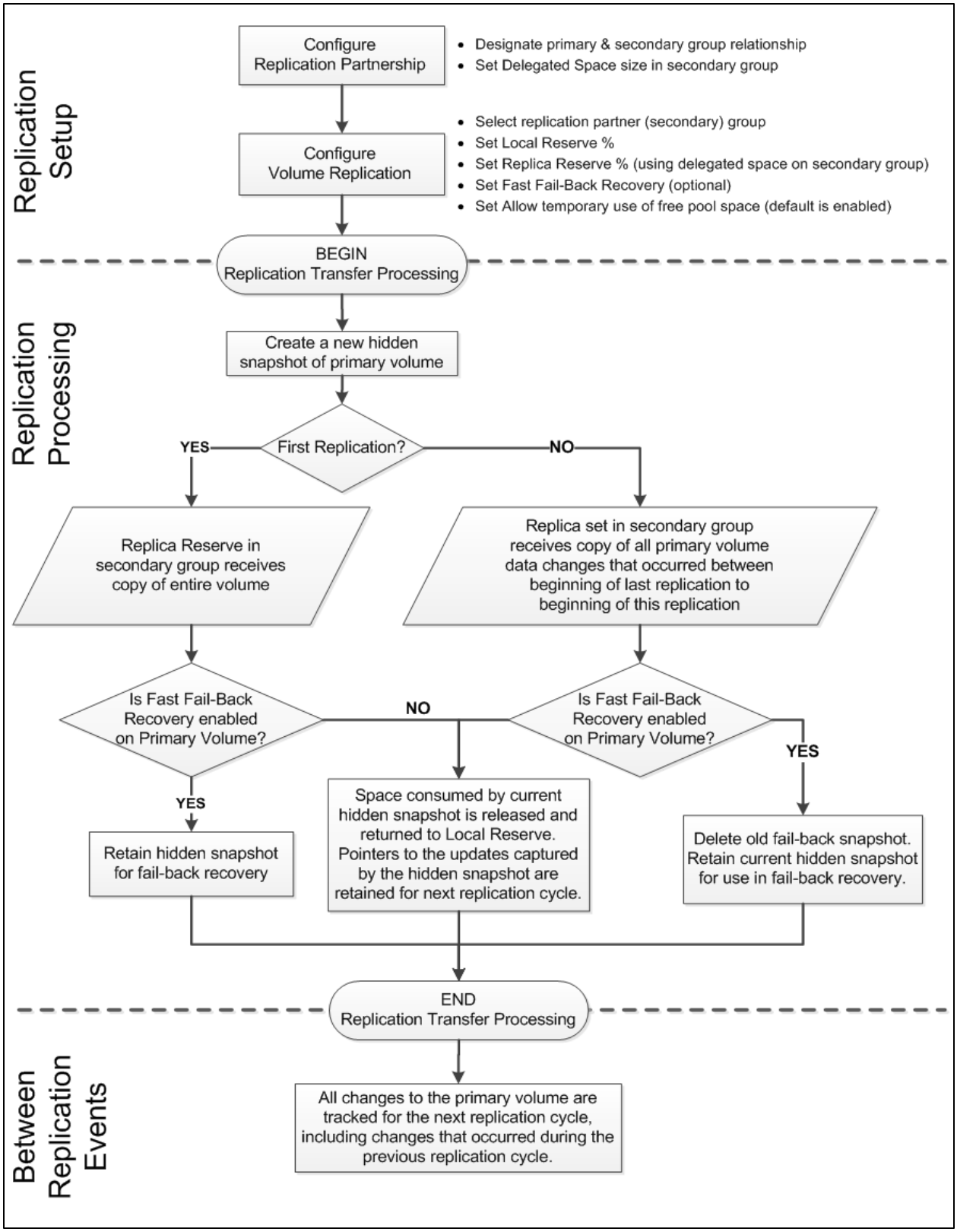

Replication Setup
Configure Replication Partnership
Designate primary & secondary group relationship
Set Delegated Space size in secondary group
Configure Volume Replication
Select replication partner (secondary) group
Set Local Reserve %
Set Replica Reserve % (using delegated space on secondary group)
Set Fast Fail-Back Recovery (optional)
Set Allow temporary use of free pool space (default is enabled)
BEGIN Replication Transfer Processing
Replication Processing
Create a new hidden snapshot of primary volume
First Replication?
YES
NO
Replica Reserve in secondary group receives copy of entire volume
Replica set in secondary group receives copy of all primary volume data changes that occurred between beginning of last replication to beginning of this replication
Is Fast Fail-Back Recovery enabled on Primary Volume?
NO
Is Fast Fail-Back Recovery enabled on Primary Volume?
YES
YES
Retain hidden snapshot for fail-back recovery
Space consumed by current hidden snapshot is released and returned to Local Reserve. Pointers to the updates captured by the hidden snapshot are retained for next replication cycle.
Delete old fail-back snapshot. Retain current hidden snapshot for use in fail-back recovery.
END Replication Transfer Processing
Between Replication Events
All changes to the primary volume are tracked for the next replication cycle, including changes that occurred during the previous replication cycle.

Figure 13 Replication Process

With replication, it does not matter if the volume is thin provisioned or uses a traditional volume since in either case only the data that has changed will be copied to the replica. On the secondary side, volumes are always thin provisioned to conserve available capacity used by the Replica Reserve for that volume.

6.1.4 Fast failback

With fast failback enabled you can ensure that the volume data preserved by the failback snapshot on the primary group always matches the volume data in the most recent replica stored on the secondary group. If you have an event that causes failover to the secondary group and the workload subsequently writes changes to the replica volume then failback snapshot supports a quicker fail back to the primary group by replicating only the changes made to the replica volume during the time it was active as a recovery volume on the secondary group. If the failback snapshot is not enabled, you must replicate the entire volume contents back to the primary group to complete the failback operation. Depending on the size of the volume, the failback scenario can take significantly longer to complete if Fast Failback is not enabled.

6.1.5 Sizing Replica Reserve and Delegated Space

The volume replication process in EqualLogic PS Series storage consumes extra storage space on both the primary and secondary group partners. In addition to the space consumed by the volume itself, each volume will require additional space in the primary group for Replication Reserve and Local Reserve, plus delegated space for storing replica sets in the secondary group. This is illustrated in Figure 14 below. A single delegated space on the secondary group must be used for all volumes received from a specific primary group.

Delegated space for a replication partnership must be assigned to a pool. Delegated space for different replication partnerships can be assigned to different pools on the secondary group.

> **Note:** Any type of application or operating system level process that causes block level data changes will affect replication data size and time to complete. For example, file system defragmentation of a large partition will significantly increase the amount of space consumed in replica reserve, delegated space on the secondary group and time to complete the next scheduled replication.

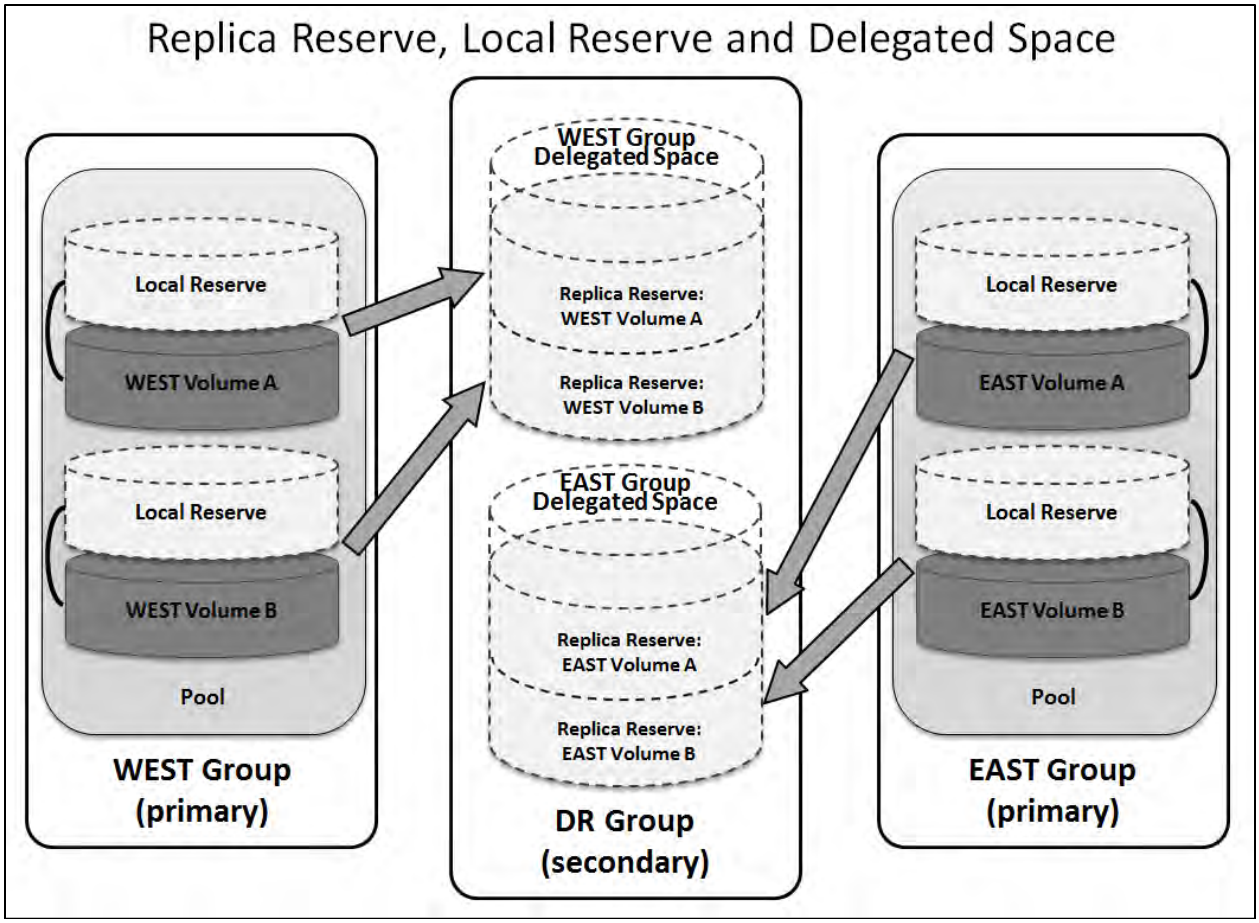


Figure 14 Replica Reserve, Local Reserve, and Delegated Space

Auto replication requires reserved disk space on both the primary and secondary groups. The amount of space required depends on several factors:

- Volume size.
- The amount of data that changes (on the source volume) between each replication period.
- The number of replicas that need to be retained on the secondary site.
- If a failback snapshot is retained on the primary group.

The default values that appear in Group Manager are sufficient to ensure that enough space is reserved for at least one successful replication, even if the entire contents of a volume are altered between replicas. Initial replication of a volume will cause the complete contents of the volume to be copied to the secondary group. Subsequent replication events will copy only the changed data. The recommended and space efficient guidelines for sizing replication reserves and delegated space are presented in Table 20.

Table 20 Replication space sizing guidelines

Replication Space	Recommended Value		Space Efficient Value
Local Reserve (Located on Primary Group.)	No Failback Snapshot:	**100%**	**5% + CR[(a)]**
	Keep Failback Snapshot:	**200%**	**10% + CR[(a)]**
Replica Reserve (Located on Secondary Group.)	**200%** Ensures there is adequate space for the last replica and any replica in progress.		**105%[(b)] + CR x (# of Replicas – 1)**
Delegated Space (Located on Secondary Group. The replica reserve space for all replica sets coming from a single group)	Must be large enough to hold the sum of all replica reserve sizes for all volumes replicating to that group.		Monitor change rate, adjust to lower than recommended value and continue monitoring.

(a) CR: "Change Rate". For details on how volume changes affect replication space, see the Centralized Replication section under Replication Configuration Options in the Group Manager Help documentation.
(b) Start with 105%, then add to that the maximum number of replicas expected to be stored in the replica set minus 1, multiplied by the expected Change Rate.

6.1.6 Effect of TCP Window Size on Latency across WAN links

In certain cases, increasing the TCP Window Size setting on PS Series controllers can improve replication throughput across WAN links. The ability to change this setting is currently not supported using the EqualLogic Group Manager. Customers interested in changing TCP Window size settings should contact Dell EqualLogic Technical Support for assistance. For more details on how this setting can affect replication performance across WAN links, see the following publication:

- **Dell EqualLogic Auto-Replication: Best Practices and Sizing Guide** : http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19854181aspx

6.1.7 Replication partner compatibility

Table 14 provides details on replication partner interoperability based on array firmware version. For the latest information, consult the release notes for your array firmware version.

Table 21 Firmware replication compatibility

Firmware of "Local" Group	Firmware on Replication Partner
V6.0	V5.1.x, V5.2.x, V6.0.x
V5.1.x, V5.2.x	V5.0.x, V5.1.x, V5.2.x
V5.0.x	V4.2.x, V4.3.x, V5.0.x, V5.1.x, V5.2.x
V4.2.x, V4.3.x	V4.1.x, V4.2.x, V4.3.x, V5.0.x
V4.1.x	V4.0.x, V4.1.x, V4.2.x, 4.3.x
V4.0.x	V3.2.x, V3.3.x, V4.0.x, V4.1.x, V4.2.x
V3.2.x, V3.3.x	V3.0.x, V3.1.x, V3.2.x, V3.3.x, V4.0.x
V3.0.x, V3.1.x	V3.0.x, V3.1.x, V3.2.x, V3.3.x

6.1.8 Clustering

To support a shared storage environment, EqualLogic allows concurrent access to any volume. Concurrent volume access is enabled on a per volume basis within the Group Manager or via CLI. The EqualLogic array will not manage concurrent access to these volumes. Control of access to a shared volume must be provided by means of access control software on each host that has access to the volume.

Warning: Failure to provide access control to a shared volume can result in data loss.

By default, EqualLogic PS Series groups disable multi-host (shared) access to target.

- If needed, you can enable multi-host access for a volume.
- If you enable multi-host access to a volume, then proper I/O operation concurrency must be maintained.

6.2 Synchronous replication

6.2.1 About Synchronous replication

Synchronous replication (SyncRep) is the simultaneous writing of volume data across two different storage pools in the same PS Series group, resulting in two hardware-independent copies of volume data. Each write must go to both pools before the write is acknowledged as complete. If one pool is not available due to a power failure or resource outage, you still can obtain volume data from the other pool.

You cannot perform traditional replication on a SyncRep-enabled volume.

6.2.2 How Synchronous replication works

SyncRep is enabled on a per-volume basis. In volumes for which SyncRep is not enabled, volume data and snapshots are located only in the pool to which the volume is assigned.

In SyncRep-enabled volumes, volume data exists simultaneously in two pools:

- SyncActive: The pool to which iSCSI initiators connect when reading and writing volume data.
- SyncAlternate: When volume data is written to the SyncActive pool, the group simultaneously writes the same data to this pool. You can switch the roles of the SyncActive and SyncAlternate pools.

When you switch the SyncActive and SyncAlternate pools, the former SyncActive pool then becomes the SyncAlternate pool, and vice-versa. No iSCSI target configuration changes are required. During the switch, host connections are logged out. iSCSI initiators can reconnect when the switch has completed. Depending on its configuration, the initiator may automatically reconnect. If you delete a volume for which SyncRep is enabled, the system will place the SyncActive volume into the Recovery Bin. However, the SyncAlternate volume will be deleted outright and cannot be recovered.

6.2.3 Synchronous states

There are three SyncRep states for a volume:

- In-Sync —SyncActive and SyncAlternate pools contain the same volume data.
- Paused—Administrator has paused SyncRep. While SyncRep is paused, the volume is still online, and initiators can connect to and write to the SyncActive volume. An administrator may pause and later resume SyncRep. For example, this could happen in a maintenance window during which the SyncAlternate pool is taken offline. If data is written to the volume while SyncRep is paused, it is written only to the SyncActive pool, and the two pools are out of sync. The group tracks all volume writes while SyncRep is paused and, when the administrator resumes SyncRep, writes the tracked changes to the SyncAlternate pool.
- Out of Sync—SyncActive pool and SyncAlternate pool may not contain the same volume data; the SyncActive pool contains the most recent volume data. A volume can become out of sync if SyncRep is paused, or if one of the pools becomes unavailable or has no free space. The volume can become out of sync when the snapshot reserve in the SyncAlternate pool is full, but only when the snapshot space recovery policy sets volumes offline when the snapshot reserve is depleted.

Whenever the volume's state changes to paused or out of sync, the group creates a snapshot of the volume that reflects the volume's contents at the point in time when the state changed. This snapshot resides in the SyncActive pool. If there is insufficient room in the snapshot reserve for the snapshot, the group does not create

6.2.4 Caveats:

Depending on the quantity of tracked changes, activity within the group, and available network bandwidth, there may be an extended period of time before the two pools become in sync again. The Group Manager GUI displays the status of this operation. When SyncRep is first enabled, or at any other time when the volume is writing data to both pools to become in sync, performance degradation may occur. This effect increases with the quantity of tracked changes, but it is significantly reduced after the volume becomes in sync.

6.2.5 Requirements for using SyncRep

Before you can configure a volume to use SyncRep, verify that the following requirements are met:

- Two pools, each containing at least one member.

- Adequate network bandwidth between pools.
- Free space in each pool to accommodate the volume and snapshot reserve for the volume.
- You cannot enable SyncRep on a volume for which traditional replication is configured, and you cannot enable traditional replication on a volume for which SyncRep is configured. See the "Disabling Replication" topic in the online help for instructions on disabling traditional replication on a volume.
- You cannot enable SyncRep on a volume that is bound to a group member.

6.2.6 How SyncRep protects volume availability

In normal SyncRep operation, in which the volume is in sync, the pools containing the SyncActive and SyncAlternate volumes contain identical data. Volume writes are accepted as shown in Fig 15.

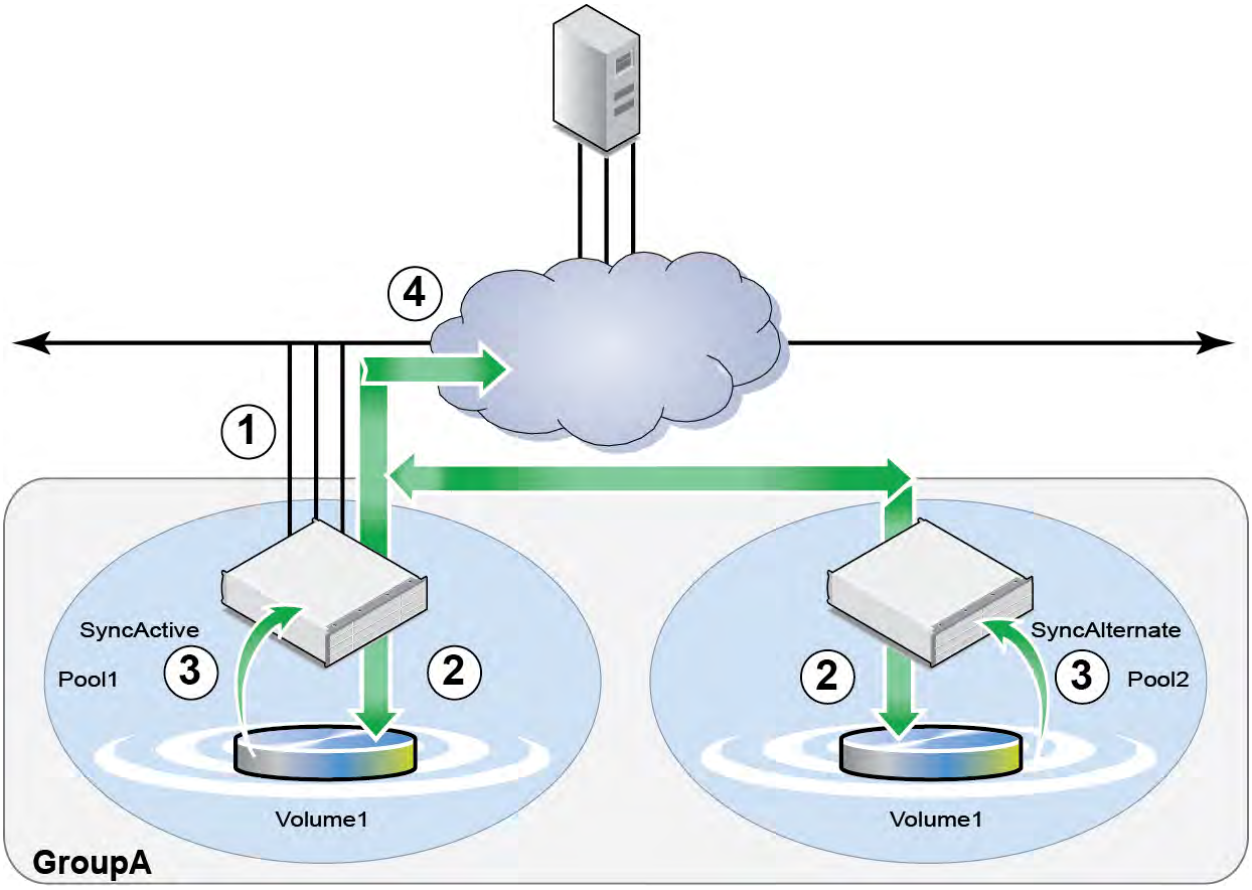


Figure 15 SyncRep

1. The iSCSI initiator sends a write to the group.
2. Data is simultaneously written to both the SyncActive and SyncAlternate volumes.
3. The SyncActive and SyncAlternate volumes confirm to the group that the writes are complete.
4. The write is confirmed to the iSCSI initiator.

6.2.7 SyncAlternate volume unavailable

If the group can write to the SyncActive volume, but can no longer write to the SyncAlternate, initiator access to the volume continues without disruption, as shown in Figure 16.

1. The SyncAlternate volume becomes unavailable.
2. Initiator access to the volume continues without interruption; the volume is out of sync.
3. The group tracks all changes written to the volume.
4. When the SyncAlternate volume becomes available, tracked changes are written to the SyncAlternate volume.

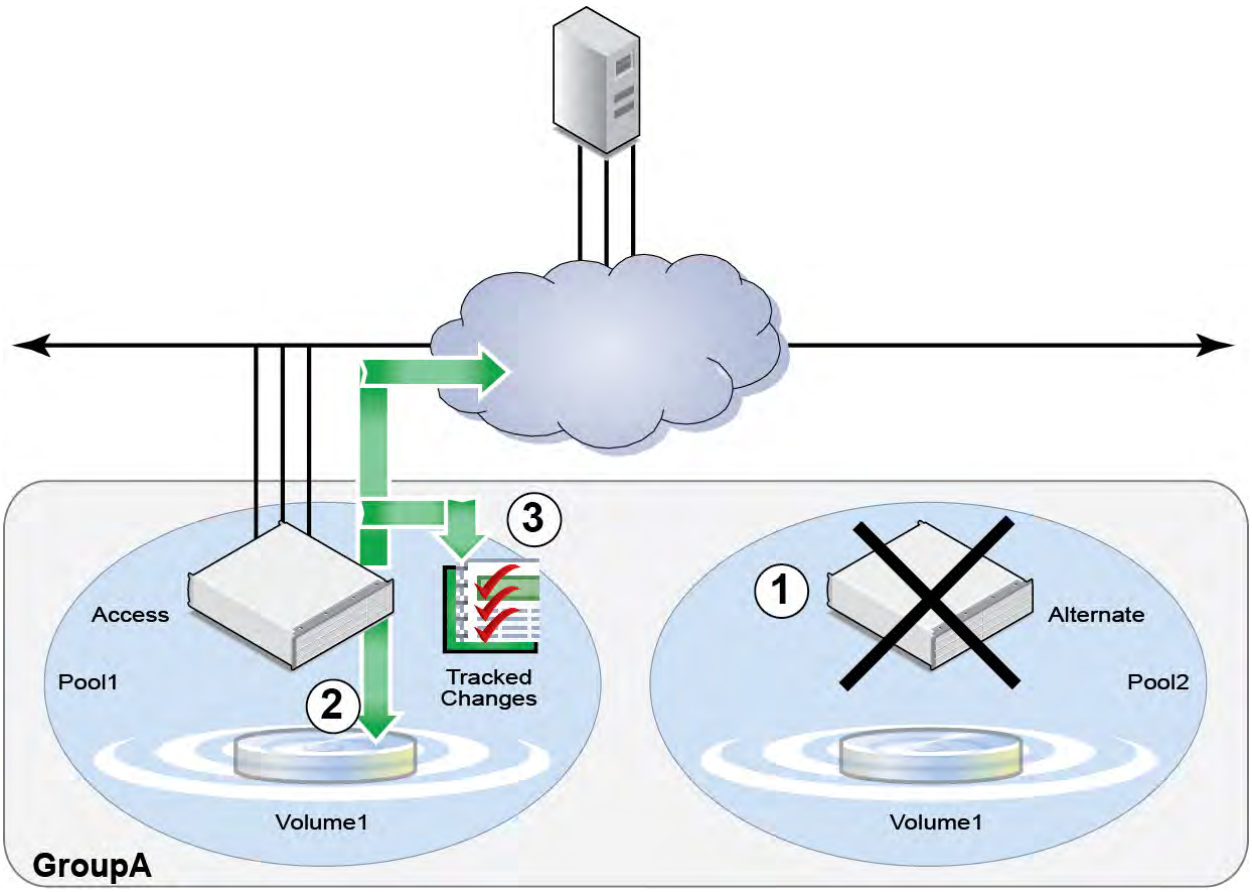


Figure 16 SyncAlternate volume unavailable

Performance will be temporarily degraded while changes are being tracked or when tracked changes are being written back to the SyncAlternate volume.

6.2.8 Tracked changes written to the SyncAlternate volume

1. The SyncAlternate volume becomes available.
2. Changes tracked while the SyncAlternate volume was unavailable are written to it. Until all tracked changes are written, the data in the SyncAlternate volume is valid up to the point in time when the volume went out of sync. During this time, performance may be temporarily degraded. When all tracked changes are written, the volume goes back in sync.
3. New writes are simultaneously written to both the SyncActive and SyncAlternate, and normal SyncRep operations resume as shown in Figure 17.

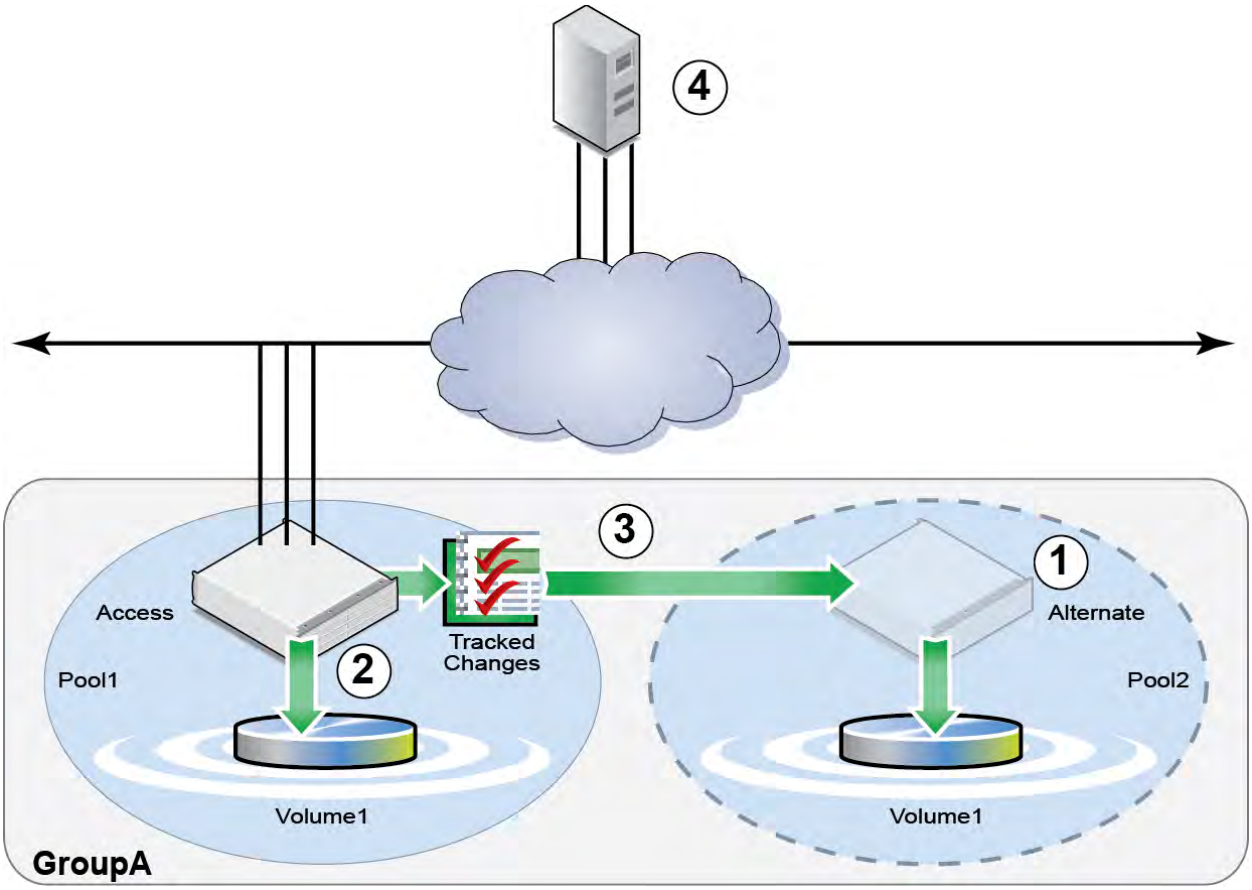


Figure 17 Tracked changes written to the SyncAlternate volume

6.2.9 SyncActive volume unavailable

If a malfunction occurs in the SyncActive pool, or some other event has occurred causing the volume to go offline, you can safely switch or failover to the SyncAlternate volume by following one of the procedures listed below.

- **Volume In Sync**: If the volume is in sync, you may switch to the SyncAlternate as documented in the online help. Although host access to the volume is disrupted during the switch, no initiator changes are required.
- **Volume Out of Sync**: If the volume is out of sync, the group administrator performs the steps outlined in the online help to safely restore access to the volume:

Table 22 below gives a comparison between traditional replication and SyncRep.

Table 22 Comparing SyncRep and traditional replication

Replication consideration	Traditional replication	SyncRep
Typical Use Case	A point-in-time process that is conducted between two Groups, often in geographically diverse locations. Replication provides protection against a regional disaster such as an earthquake or hurricane. Traditional replication has the advantage of providing multiple recovery points. A disadvantage of traditional replication is that the state of the data between recovery points is unknown; if any changes are made to the volume since the last replica was created, they could be lost.	A real-time process that keeps two identical copies of volume data in two different pools within the same PS Series group. SyncRep is useful for maintaining two copies of a volume's data in the same data center, or dispersed to two different facilities on the same campus or in the same metropolitan area. An advantage of SyncRep is that it captures a duplicate copy of every write. One disadvantage of this is that if an application writes bad data to the volume, the bad data is simultaneously written to both the SyncActive and SyncAlternate pools.
Recovery Time	If a disaster occurs in the primary group, you can promote the replica set on the secondary group to a recovery volume. After the promotion, you must reconfigure initiators to discover and log in to the iSCSI target now hosted by the secondary group, or switch to an alternate set of server resources that have been preconfigured to use the secondary group storage. See the Impact on Applications row of this table for more information.	If a disaster involving the SyncActive pool occurs, you can manually switch the volume to the SyncAlternate pool. After the switch, the SyncAlternate pool becomes the SyncActive pool and hosts the volume. Host access to the volume is disrupted by the switch, but iSCSI initiators do not need to be reconfigured.
Recovery Point	The recovery volume contains point-in-time data that is current as of the most recent replica. Replication can be scheduled to take place as frequently as once every five minutes. You can also restore to the point in time when any previous replicas were created, provided that the replicas have been retained.	SyncRep provides a single recovery point: the most recent acknowledged write to the volume.

Replication consideration	Traditional replication	SyncRep
Network Requirements	Replication requires that the network connection between the primary and secondary groups must be able to handle the load of the data transfer and complete the replication in a timely manner.	Because writes are not acknowledged until they are written to both the SyncActive and SyncAlternate pools, SyncRep is sensitive to network latency. The network must be able to handle the load of the data transfer from the SyncActive pool to the SyncAlternate pool and complete the replication in a timely manner, or application performance may suffer.
Snapshots	Replication is functionally similar to snapshots, creating point-in-time copies of the volume. If the "keep failback" option is enabled, the group creates a "failback snapshot" on the primary group every time a replica is created. This allows for "fast failback" capabilities. In addition, you can schedule the creation of snapshots, or create them on demand, just as you would with any other volume. See About SyncRep and Snapshots on page 126 for more information.	SyncRep creates snapshots of the volume whenever the SyncActive and SyncAlternate pools are switched.
Scheduling	Replication operations can be scheduled using the same mechanism used for scheduling snapshots.	Replication between the SyncActive and SyncAlternate pools is continuous. Therefore there is no need or mechanism to schedule SyncRep.
Pool Space Requirements	The primary group must have enough space for the volume reserve and local replication reserve, in addition to any snapshot reserve. The secondary group must have enough free space delegated to the primary group for the volume reserve and the replicas that record changes to the volume's data over time.	Both the SyncActive pool and the SyncAlternate pool must have enough space for the volume and snapshot reserve.

Replication consideration	Traditional replication	SyncRep
Impact on Applications	iSCSI initiators must be reconfigured to connect to the secondary group after the failover, or an alternate set of host resources must be brought online, both of which may cause application disruptions. If you are using the Host Integration Tools, you can coordinate replication with host software to quiesce applications on a schedule and create application consistent Smart Copies. Replication can help protect against the corruption of application data: depending on when the replica occurred and what your replica retention policies are, you may be able to restore the volume to a point in time before the corruption occurred.	Pool switches may cause disruptions in host access to the volume, but no change to the iSCSI initiator configuration is required to restore access. Writes must be committed to both pools before they are acknowledged to the host, so the application must be able to tolerate whatever additional delay is caused by the simultaneous writes. When SyncRep is first enabled, or at any other time when the volume is writing data to both pools to become in sync, performance degradation may occur. This effect is diminished after the volume becomes in sync.
PS Series Group Requirements	Two PS Series groups, each of which must contain at least one member.	One PS Series group containing two storage pools, each of which must contain at least one member.

6.3 Protecting your EqualLogic group with Internet Protocol Security

Internet Protocol Security (IPsec) is a set of standardized protocols designed to allow systems on IP-based networks to verify each other's identities and create secured communication links. IPsec uses cryptographic security mechanisms for authentication and protection. IPsec validates the identity of devices communicating over IP-based networks, encrypts all data passing between participating systems, and protects against disclosure, modification, eavesdropping, and attack. IPsec is supported for both IPv4 and IPv6 networks.

In the context of an iSCSI SAN that uses EqualLogic PS Series storage arrays, IPsec secures communications between group member arrays and also between iSCSI initiators and the group. You can use policies to configure your IPsec implementation to protect iSCSI traffic based on initiator IP address, initiators in a specific subnet, or network protocol. IPsec authentication is handled using certificates or pre-shared keys.

IPsec is supported only for PS Series array models PS6xxx, PS41x0, and PS-M4110, and can only be enabled for a group if all members support IPsec. See the Dell EqualLogic PS Series Storage Arrays Release Notes for more information.

6.3.1 Types of protected traffic

The types of traffic protected by IPsec are shown in Figure 18 and the sections that follow. Any incoming or outgoing IP traffic that travels between hosts and the group can be protected with IPsec.

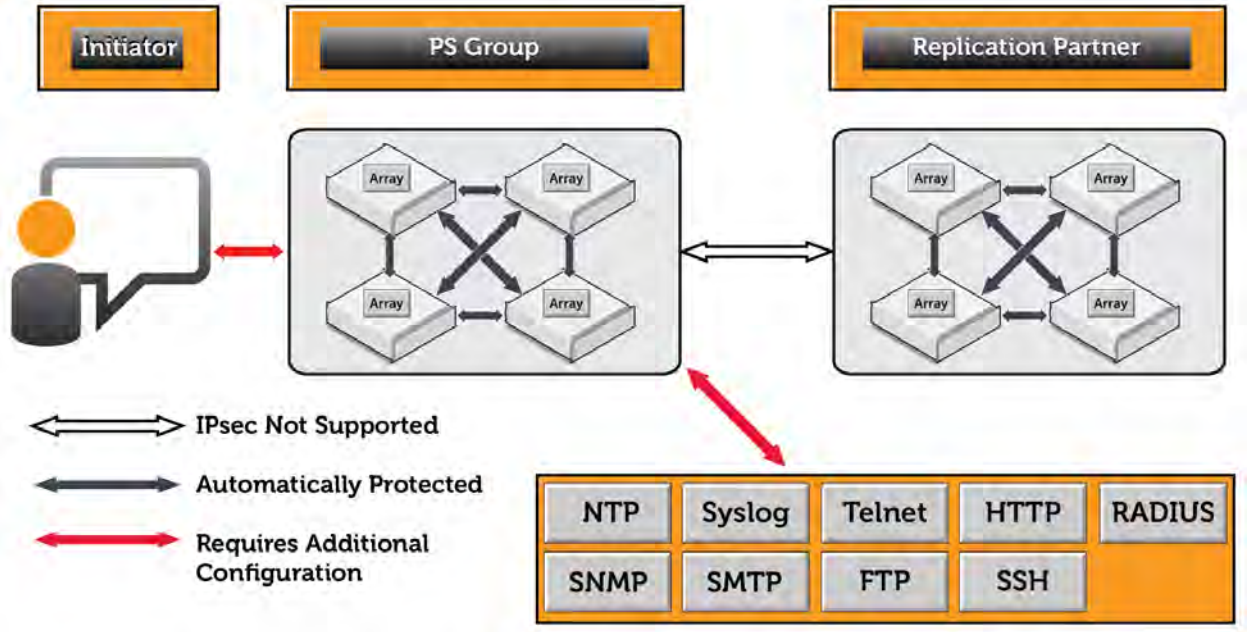


Figure 18 IPsec protected traffic

6.3.2 Protected Intra-Group Traffic

Once IPsec is enabled, all network traffic between group members is automatically protected with IPsec using IKEv2. No further configuration is required.

6.3.3 IPsec and Replication

The PS Series Firmware provides no mechanism for using IPsec to protect traffic between replication partners. It is technically possible to create IPsec polices on both the primary and secondary group in which each group treats the other as an iSCSI initiator and traffic is protected accordingly. However, this is an unsupported configuration, and Dell recommends against implementing it in a production environment.

6.3.4 About IPsec Security Parameters

IPsec security parameters control the authentication and key negotiation carried out using the Internet Key Exchange IKEv1 or IKEv2 protocol.
Security parameters specify the following:
Using IKEv1, IKEv2, or manual keying.
You can configure IPsec to use manual keys. However, manual keying provides significantly weaker security than IKEv1 or IKEv2, and is also significantly more difficult to configure. Consequently, Dell strongly discourages the use of manual keying in any production environment. IKEv1 or IKEv2 are the preferred keying methods. Refer to the Dell EqualLogic Group Manager CLI Reference Guide for more information about using manual keys.

6.3.5 About IPsec Certificates

Certificates are used in an IPsec configuration as one method of authenticating secured connections between iSCSI initiators and the group. Implementation of an IPsec-enabled SAN requires both a root-CA (Certificate Authority) certificate from the issuing authority and a local certificate to authenticate the group.

You can generate certificates suitable for use in IPsec connections to the PS Series using any Windows, OpenSSL, or other commercial Certificate Authority product.
You can use the Group Manager CLI to import, display, and delete certificates. using the IPsec certificate commands. See the Dell EqualLogic Group Manager CLI Reference Guide for more information.

6.3.6 About IPsec Pre-Shared Keys

In addition to using certificates, you can use pre-shared keys to authenticate secured connections. Pre-shared keys are identical strings that are specified at both ends of the communications pathway. This allows the systems to correctly identify each other.
You can use either ASCII or hexadecimal strings. ASCII can be used in most situations. However, you can also use hexadecimal strings if your organization mandates their use, if you have systems that do not support the use of ASCII strings, or if you want to use unsupported characters.

6.3.7 About IPsec policies

Traffic that meets the conditions stipulated by the policy can either be passed, dropped, or protected using an IPsec security parameter associated with the policy.
You can use IPsec policies to apply IPsec protection to traffic that meets one or more of the following criteria:

- Data traveling to or from specific IP addresses, or a range of IP addresses defined by a specific subnet or netmask
- IPv4 or IPv6 traffic
- Specific network protocols: TCP, UDP, or ICMP (either IPv4 or IPv6)

Unless explicitly specified by the policy, traffic is allowed to pass. If you want to drop all traffic that is not explicitly protected or passed, you must create an IPsec policy that drops traffic by default.
If there are multiple IPsec policies in place, the system determines their priority by the order in which they were created; policies created first take precedence over policies created later.
You can also use IPsec policies to determine what traffic is being protected using IPsec, and what traffic is being passed or dropped without encryption.
IPsec policies are managed using the IPsec policy commands. See the Dell EqualLogic Group Manager CLI Reference Guide for more information.

6.3.8 IPsec considerations and limitations

The limitations listed in the sections below apply when implementing IPsec.
Configuration limitations

- IPsec is only supported for certain PS Series array models, and can only be enabled for a group if all members support IPsec. See the Dell EqualLogic PS Series Storage Arrays Release Notes for more information.
- IPsec can only be enabled and configured using the Group Manager CLI. The Group Manager GUI provides no facility for configuring or monitoring IPsec.
- The PS Series array does not serve as an IPsec-secured gateway; it only behaves as an IPsec-secured host.
- You cannot use the save-config CLI command to preserve the group's IPsec certificates and pre-shared keys. The save-config command saves the CLI commands that were used to configure IPsec, but it does not save certificates that have been transferred to the array using FTP. Therefore, when you restore a configuration, you must manually restore any configuration options set using the IPsec certificate load, IPsec security-params create certificate, and IPsec security-params pre-shared-key commands.

- Kerberos-based authentication is not supported.
- Multiple Root Certificate Authorities (CA) are not supported.
- Certificate Revocation Lists (CRL) are not supported.
- Only users with group administrator privileges can configure IPsec.
- Perfect Forward Secrecy (PFS) is not supported.
- Encrypted private keys are not supported for X.509 format certificates.
- Dell recommends using a minimum of 3600 seconds and 10GB lifetime rekey values.
- IKE mobillity is not supported
- NAT Traversal (NAT-T) is not supported. Dell recommends against placing a firewall that performs address translation between the PS Series group and its IPsec peers.
- If you use the Windows default IPsec lifetime rekey values, the high rekey rates may be disruptive for protected iSCSI traffic. Values in the range of 1GB to 100GB, depending on iSCSI traffic, are recommended instead.

6.3.9 Performance considerations

The performance impact of IPsec varies by host and network configuration, and increases with the number of IPsec-protected iSCSI connections to the group. Even if IPsec is only used to protect traffic between group members, I/O performance is still affected. Based on these factors, you can expect that using IPsec may degrade I/O performance.
Although PS Series group members use hardware to accelerate cryptographic operations, many initiators perform these operations in software, which can cause a further reduction in the speed of communications between iSCSI initiators and the group.

6.3.10 Host Connectivity Considerations

Enabling or disabling IPsec for the group using the IPsec enable and IPsec disable commands might disrupt host connectivity to the group for several minutes. To prevent unplanned outages, IPsec should be enabled or disabled during a planned maintenance window when there are no active iSCSI connections to any volumes.

Consult the documentation for your host operating systems, HBAs, and iSCSI initiators to verify that they support IPsec. There might also be known issues and idiosyncrasies with the initiators' IPsec support that require additional planning or configuration.

When configuring IPsec with Windows hosts, note the following:
IPsec traffic is not always handled correctly if the IPsec policy is configured to protect only a subset of traffic between the host and the group. For example, if the IPsec policy protects only iSCSI traffic on port 3260, the Windows host may not perform reliably when connecting to the group. As a workaround, IPsec policies should apply to all traffic passing between the group and Windows systems. Microsoft KB article 2665206 discusses this in greater detail.

IPsec must be configured using the Windows Firewall with Advanced Security. Do not use the IPsec option in the Microsoft iSCSI initiator, which does not have the capability to fully configure an IPsec configuration between the host and the group. Further, if you attempt to configure an IPsec connection using the iSCSI initiator, the system might not allow you to remove the partial configuration and replace it with a complete configuration created with Windows Firewall.
IPsec policies defined using the Local Security Policy Manager are not supported.

7 EqualLogic SAN design

An EqualLogic iSCSI SAN can be operated in any network that supports the industry standards and IP subnet design guidelines described in this section. Because of this flexibility, there are many network design and configuration choices that can affect SAN performance. The following sections provide details related to network design and configuration to support the use of an EqualLogic SAN.

> **Note:** With the addition of FS Series NAS functionality, the EqualLogic product family now provides an iSCSI based **unified file and block storage platform**. An EqualLogic SAN can now be used to provide **block level access** (direct iSCSI access to PS Series arrays) or **file system level access** (via the FS Series appliance) using NFS or CIFS protocols and the Dell FluidFS scale-out file system.

Unless otherwise stated, recommendations in this document are applicable to both file and block environments.

7.1 General requirements

7.1.1 Implementation of standards

EqualLogic SANs are based on industry standards. The following standards are required to support all host to target communications and member to member communications in an EqualLogic SAN:

- IETF Standards
  - IETF RFC1122 "Requirements for Internet Hosts – Communications Layers"
  - IETF RFC1123 "Requirements for Internet Hosts – Application and Support"
  - IETF RFC3270 "Internet Small Computer Systems Interface (iSCSI)"
- IEEE Standards
  - 802.1
  - 802.3

iSNS support

An Internet Storage Name Service[1] (iSNS) server can support discovery, management and configuration of group members by dynamically updating information about the iSCSI target names for group volumes. Once the IP address of an iSNS server is entered in an iSCSI initiator's configuration utility, the setting is persistent across initiator sessions. A PS Series group can be configured to register with up to three iSNS servers.

> **Note:** Starting with Firmware V4.1.4, volume and snapshot identifiers are no longer automatically published to iSNS servers. This applies to new volumes and snapshots as well as volumes and snapshots that existed before the group was upgraded to V4.1.4.

[1] The Internet Storage Name Service (iSNS) specification: http://tools.ietf.org/html/rfc4171

7.1.2 General requirements and recommendations

For EqualLogic PS Arrays, the following general SAN design requirements apply:

- For all members (arrays) in a given SAN Group all ports should be connected to the same subnet. This allows the arrays to communicate with each other as a group of peer members. The arrays must be in the same subnet as the group's "well known" IP address.

> **Note:** Hosts can be in a different subnet as long as those hosts have layer 3 routing available to the subnet containing the arrays and the group's well known IP address.

- It is strongly recommended that a physically separated network be used for iSCSI traffic and that this network not be shared with other traffic types.

> **Note:** If there is a requirement to share the same physical networking infrastructure with other non-iSCSI traffic then Data Center Bridging (DCB) is the recommended method for sharing networking resources.

- Rapid Spanning Tree Protocol must be enabled if the SAN infrastructure has more than two switches in a non-stacked configuration, and portfast must be enabled on all edge device ports (hosts, FS Series appliances and arrays).
- Port density requirements to support fully redundant configurations and maximum SAN throughput are as follows:
  - **PS4x00 family**: 2x 1GbE ports per controller = 4x 1GbE ports total
  - **PS4110 family**: 1x 10GbE port per controller = 2x 10GbE ports total
  - **PS5x00 family**: 3x 1GbE ports per controller = 6x 1GbE ports total
  - **PS6x00 family**: 4x 1GbE ports per controller = 8x 1GbE ports total
  - **PS6510 family**: 2x 10GbE ports per controller = 4x 10GbE ports total
  - **PS6110 family**: 1x 10GbE port per controller = 2x 10GbE ports total
  - **FS7500 NAS**: 12x1GbE ports per controller node (four client LAN and eight iSCSI SAN) + 1 100/1000Mb port per controller node for IPMI interconnection
  - **FS7600 NAS Applicance: 16x1GbE Ethernet ports per appliance = 8 x 1Gb Ethernet ports per NAS controller for client connectivity and 8 x 1GbE Ethernet ports per NAS controller for SAN connectivity**
  - **FS7610 NAS Appliance: 8x10GbE Ethernet ports per appliance = 4 x 10GbE SFP+ Ethernet ports per NAS controller for client connectivity and 4 x 10GbE SFP+ Ethernet ports per NAS controller for SAN connectivity**
- At least two iSCSI SAN ports per host (block level iSCSI access) are required for fully redundant SAN connectivity. Host ports can be 1GbE or 10GbE.
- Quality of Service (QoS) based on what is traditionally designated as IEEE 802.1p is not currently supported for use with EqualLogic SANs. QoS and Class of Service designations must be disabled.
- All switches within the SAN must be interconnected such that there is always a path from any Ethernet port on one array to all other Ethernet ports on all other arrays in the group.
- All switches and host network controllers within the infrastructure must have flow control enabled for optimal performance.

- Any EqualLogic SAN group that is required to send or receive replication traffic to/from another SAN group must have an uninterrupted communications path (ie. "visibility") between each group.
- To prevent a switch failure from also disabling all paths between a host and its connected volumes, all ports from each controller need to be connected to at least two different switches
- For PS4100/PS6100 family arrays, split the vertical port pair connections between two switches to ensure 100% bandwidth capability is maintained in the event of a vertical port failover event. See the configuration diagrams in Section 7.3.6.
- The above guideline regarding path redundancy across multiple switches also applies in FS Series NAS appliance configurations. (Refer to the connection diagrams in Section 13.1)
- Do not block IPv6 traffic on the SAN internal switches when utilizing FS Applicances
  FS76x0 uses IPv6 for internal communication and setup only; not for user data.
  Client communication happens over IPv4
- **For SANs connecting to an FS Series appliance, all switches in the SAN must have jumbo frames enabled.**

For EqualLogic PS Series Arrays, the following general SAN design recommendations apply:

- Take advantage of your switch's VLAN capabilities. You should create a VLAN dedicated to iSCSI traffic (even on dedicated switches). If necessary, create a second VLAN for management traffic. The actual VLAN configuration of your iSCSI SAN will be dictated by your SAN network design requirements and the features of the iSCSI SAN switches being used.
- Jumbo frames should be enabled. If you choose to use jumbo frames then all nodes in the SAN fabric must have jumbo frames enabled.
- For best performance and reliability, we recommend that all interconnection paths between non-stacking switches (LAGs) use a dynamic link aggregation protocol such as LACP

7.1.3 Quality of service (qos)

Quality of service is described as either of the following:

- The ability to provide different priority levels to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.
- A network function implemented in some routers and switches that provides high priority for certain types of latency-sensitive traffic (for example, VoIP) and lower priority for other types of traffic (for example, web or http).

PS Series arrays are designed to provide I/O as fast as your network can support it. Therefore, using QoS with iSCSI traffic does not produce expected or desirable results on the SAN. Also, QoS rules can affect how well–or even whether–replication succeeds between PS Series groups. If you plan to use QoS, Dell recommends that you use it only on VLANs that do not carry iSCSI traffic, or on WANs, where bandwidth is shared with other applications and the PS Series array uses it for time-insensitive replication traffic.

Dell recommends against using QoS on the SAN.

7.2 Ethernet switches and infrastructure

Any switch used in an EqualLogic SAN should meet the requirements listed in this section.

Note: A detailed and frequently updated list of recommended switches is maintained in a separate document:
Validated Components List for EqualLogic PS Series SANs
http://www.delltechcenter.com/page/EqualLogic+Validated+Components

Note: The FS7500 NAS appliance requires the use of 1Gb switches that meet the requirements in this section.

An EqualLogic SAN consists of one or more hosts connected to one or more PS Series arrays through a switched Ethernet network.

Note: The minimum network configuration for a PS Series Array consists of a connection between Ethernet0 on each control module and a network switch. To increase performance and availability, configure multiple network interfaces on an array and connect them to multiple switches. EqualLogic does not support Direct Attached Storage (DAS) configurations.

To support a high performance Dell EqualLogic SAN, switches must meet the following general requirements:

- **Low latency:** Switches with relatively high latency may cause SAN throughput performance to degrade, and under high load conditions they could increase the risk of dropped connections.
- **Non-blocking backplane design:** SAN Switches should be able to provide the same amount of backplane bandwidth to support full duplex communication on ALL ports simultaneously.
- **Adequate buffer space per switch port:** In addition to supporting data transfers between the hosts and the SAN, Dell EqualLogic arrays also use the SAN to support inter-array communication and data load balancing. For this reason, the more buffer space per port that a switch can provide the better. Due to the multitude of buffer implementations used by switch vendors, Dell cannot provide definitive guidelines as to how much is enough. Port buffers should be designed such that data is not lost when traffic reaches extreme levels. Due to the clustered storage traffic patterns used by the EqualLogic SAN architecture, switches that support "cut-through" mode are not suitable for use in an EqualLogic SAN and may actually result in lower overall SAN performance.
- **Support for IEEE 802.3x flow control (passive and/or active) on ALL ports:** Switches and network interface controllers used in an EqualLogic SAN must be able to passively respond to any "pause" frames received. If possible, you should use switches that have the ability to transmit "pause" frames to external devices in the event that the device cannot adequately forward traffic in a timely fashion.
- **Support for Jumbo Frames:** This is not a requirement. But, the use of jumbo frames may yield desirable results. Most iSCSI SAN implementations should benefit from using jumbo frames. The actual impact on SAN throughput when using jumbo frames will depend on your workload's I/O characteristics.
- **Support for Rapid Spanning Tree protocol (IEEE 802.1w), or Cisco "portfast" functionality if the SAN infrastructure will consist of more than two switches:** For SAN infrastructures consisting of

more than 2 non-stacking switches, R-STP must be enabled on all ports used for inter-switch trunks. All non-inter-switch trunk ports should be marked as "edge" ports or set to "portfast".

- **Support for unicast storm control:** iSCSI in general, and Dell EqualLogic SANs in particular can send packets in a very "bursty" profile that many switches could misdiagnose as a virally induced packet storm. Since the SAN should be isolated from general Ethernet traffic, the possibility of actual viral packet storms occurring is non-existent. In an EqualLogic SAN, the switches must always pass Ethernet packets regardless of traffic patterns.
- **Support for Inter-Switch Trunking (IST) or Stacking:** IST support is required to link all switches in SAN infrastructure together. For stacking capable switches, the use of stacking ports for IST is assumed. A good rule of thumb for stacking link bandwidth would be a minimum 20 Gbps full-duplex.
- **Support for VLAN functionality.**

> **Note:** VLANs can be used as a means of isolating out-of-band EqualLogic management traffic. Dell recommends this strategy where networking resources are limited. Otherwise, it is recommended to use a physically separated network dedicated to iSCSI traffic that is not shared with other traffic. If sharing the same physical networking infrastructure is required, then use Data Center Bridging (DCB) for EqualLogic SAN.

- **Support for creating Link Aggregation Groups (LAG):** For non-stacking switches, the ability to bind multiple physical ports into a single logical link for use as an inter-switch trunk (IST) is required. The switch should support designating one or more ports for IST (via Link Aggregation Groups). The switch should support creation of LAGs consisting of at least eight 1Gbps ports or at least two 10Gbps ports.

> **Note:** For 1GbE SANs, using non-stacking switches to connect three or more EqualLogic arrays into a single group may negatively impact SAN I/O throughput performance.

7.2.1 Connecting SAN switches in a Layer 2 network

When more than one SAN switch is required, each switch connected to the array group members will be in the same subnet. These switches must be interconnected to provide a single switched Ethernet fabric. Figure 19 shows the two common methods for interconnecting switches, using either stacking switches or non-stacking switches.

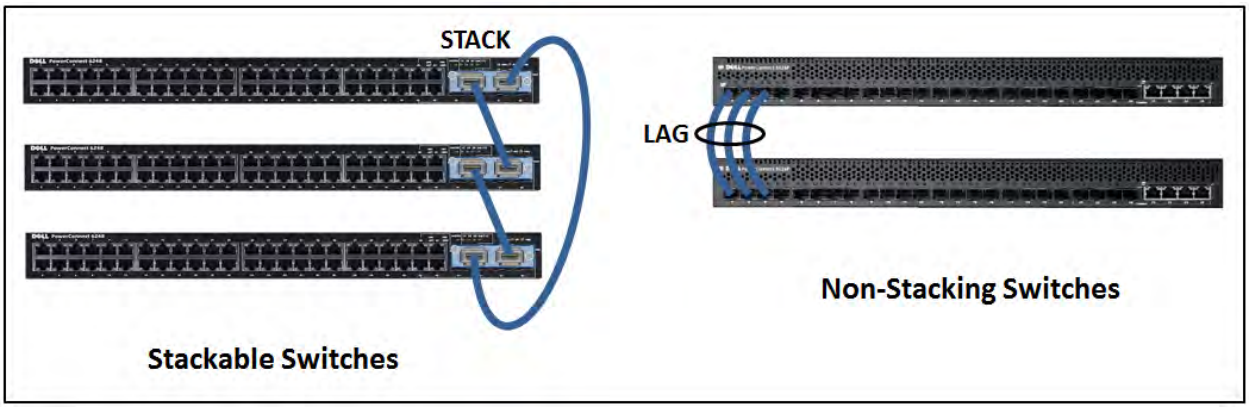


Figure 19 Switch Interconnects

7.2.1.1 Stacking Switches

Stacking switches provide the preferred method for creating an inter-switch connection within a Layer 2 network infrastructure. Stacking is typically accomplished using a vendor proprietary, high-bandwidth, low-latency interconnect that allows two or more switches to be connected in such a way that each switch becomes part of a larger, virtual switch. A stackable switch will provide a set of dedicated stacking ports. Installation of an optional stacking module may be required. Considerations for stacking link bandwidth:

1Gb Switches The stacking link bandwidth should be at least 10Gb/s in each direction on each wire (full-duplex) to provide adequate throughput to support an EqualLogic SAN consisting of 1Gb arrays.

10Gb Switches The stacking link bandwidth should be at least 40Gb/s in each direction on each wire (full-duplex) to provide adequate throughput to support an EqualLogic SAN consisting of 10Gb arrays or a mix of 1Gb and 10Gb arrays.

7.2.1.2 Non-Stacking Switches

Non-stacking switches do not have a dedicated switch interconnect. In this case to create an interconnect between the switches you must utilize one or more ports on each switch to create a Link Aggregation Group (LAG). This type of Inter-switch connection should utilize link aggregation functions (if provided by the switch) to aggregate multiple Ethernet ports into a single, logical high bandwidth interconnect. There are several options depending on the vendor of the switch. Table 23 describes the most common options.

Note: For best performance and reliability, we recommend that all interconnection paths between non-stacking switches (LAGs) use a dynamic link aggregation protocol such as LACP.

Table 23 Link aggregation types

Link aggregation type	Notes
Static	Static link aggregation defines a set of links that provide a point to point connection between two switches. These links may or may not provide failover redundancy or traffic load management.
LACP	Link Aggregation Control Protocol is based on IEEE 802.3ad or IEEE 802.1AX. LACP is a dynamic LAG technology that automatically adjusts to the appearance or disappearance of links within the defined LACP group.
PAgP	Port Aggregation Protocol (PAgP) is a Cisco Systems® proprietary networking protocol, which is used for the automated, logical aggregation of Ethernet switch ports, known as an etherchannel. This means it can only be used between Cisco switches and/or switches from licensed vendors.
Vendor Proprietary	Several switch vendors may provide additional link aggregation options that are completely proprietary or may be extensions to one of the two previously mentioned solutions. In most cases, this type of link aggregation solution is designed to reduce or eliminate the need – and the overhead – of the Spanning Tree Protocol that must be used in the two previous options. If available, these proprietary options should be considered. They may be very useful in allowing the network administrator to create a more efficient multi-switch layer 2 network infrastructure for a SAN. Be aware that these proprietary solutions must support the transmission of IEEE 802.1x flow control and jumbo frames (if used) to properly support an EqualLogic SAN.

7.2.1.3 Using a LAG to Connect Stacked Switches

In some situations it may become necessary to expand the EqualLogic SAN by using more than one single switch stack. For example, you can link multiple stacks by creating a multi-link LACP based LAG between the switch stacks. A simplified stack plus LAG switch configuration is illustrated in Figure 20. You should consider the following recommendations when designing this type of SAN:

- If possible, use 10Gb connections for all links between each stack.
- Distribute the links across multiple switches within each stack (this is known as a "cross-stack" link aggregation group).
- Use LACP or another type of dynamic link aggregation protocol.
- Perform tests to determine the best hashing algorithm to use within the link aggregation group (port channel).

Note: A multi-stack SAN infrastructure as described in this section may provide added reliability to the SAN environment. But, it may also introduce additional latency and the potential for lower throughput. The SAN designer will have to carefully consider the performance and reliability implications.

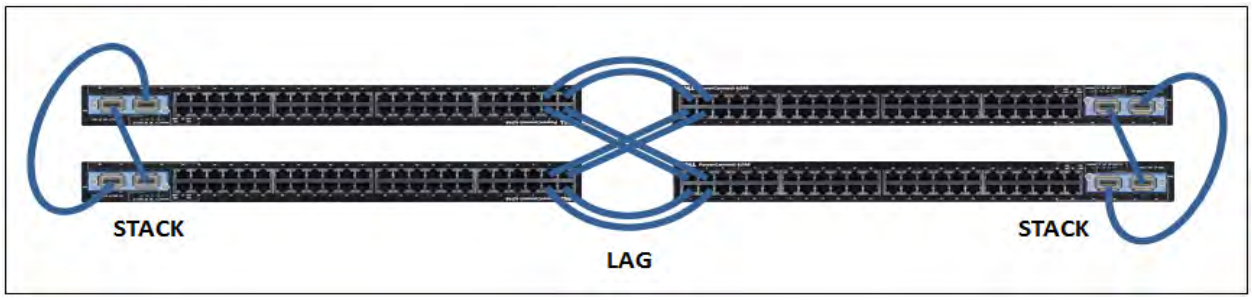


Figure 20 Using a LAG to Interconnect Switch Stacks

7.2.2 Sizing inter-switch connections

Use the guidelines in Table 24 as a starting point for estimating Inter-switch connection sizes.

Table 24 Switch Interconnect Design Guidelines

Connection Speeds	Interconnection Guidelines
1GbE switches attached to 1GbE array controllers	**1-5 arrays:** 1Gb of IST bandwidth per active array controller port (up to the aggregated maximum bandwidth of the IST. **6+ arrays:** Use 1-5 array rule, then add 1Gb of additional bandwidth for each array added
10GbE switches attached to 10GbE array controllers	**PS6010/PS6510 or PS4110/6110 (Random Small Block Workloads):** **1-5 arrays**: 20 – 30Gb of IST bandwidth between each switch **6+ arrays:** At least 6Gb of IST bandwidth per array between each switch **PS4110/PS6110 (Large Block Sequential Workloads):** **1-5 arrays:** 10Gb of bandwidth per active array controller port (up to the aggregated maximum bandwidth of the IST). **6+ arrays:** Use 1-5 array rule, then add 10Gb of additional bandwidth for each array added
1GbE switches connecting to 10GbE switches in a mixed speed SAN	Actual requirements will vary. For more details see Section 8.

7.2.3 Comparing inter-switch connection types

Table 25 provides details and recommendations for selecting interconnect options.

Table 25 Stacking vs. Inter-Switch Trunking

Interconnect type	Primary purpose	Analysis
Stacking	Create a larger, logical switch within an isolated physical location.	Advantages: Easier to manage multiple switches as single switch Higher bandwidth than using link aggregation and Ethernet Not limited by Ethernet standards Concerns: Proprietary, cannot be used to interconnect switches from different vendors Increases cost of switch Stack bandwidth overload risk Recommendation: Best way to scale a storage network in a single location on a single subnet Provides lower latency and higher bandwidth than inter-switch trunking Understand the stacking technology and limit stack size in accordance with total throughput requirements and connection counts.
Link Aggregation Groups (LAG)	Create a data path between switches in one location or subnet with those in another location or subnet	Advantages: Leverages Ethernet standard extensions Can be used to interconnect switches from different vendors Can use Link Aggregation Protocols (LACP/EtherChannel) to pool multiple 1GbE or 10GbE links into a single logical link providing bandwidth and redundancy Concerns: Most solutions limited to 8 port link aggregation group Spanning Tree Protocol must be used if more than two switches are used causing some links to be "blocked" reducing bandwidth availability LAG bandwidth overload risk Recommendation: Use when stacking is not available Use when connecting to aggregation/core switching infrastructure Use when switches are from different vendors

7.3 Building a high-availability SAN

Designing a redundant SAN requires the availability of redundant NICs or HBAs on each server. A redundant NIC configuration on the server requires at least two NICs installed into separate PCI slots in the server. Table 26 below shows how to achieve redundant server NIC connection configurations for a server with three installed NICs.

Table 26 Redundant Server NIC Configurations

	LOM NIC	Installed NIC 1	Installed NIC 2
NIC Connections to SAN	X	X	-
	-	X	X
	X	-	X

7.3.1 Design guidelines for host connectivity in a redundant SAN

Using the Dell PowerEdge R610 server as an example, you configure redundant connection paths to the SAN switches as shown in Figure 21 below. The R610 server shown in Figure 21 has one additional dual-port PCI-E NIC installed. This configuration leaves two unused ports on the LOM controller for connecting to the server LAN.

Note: As a best practice, we recommend using the same NIC devices in your server for all connection paths to the iSCSI SAN. This will minimize the complexity of server configuration management.
For maximum performance, ensure that the PCI express slot hosting the network controller has the same specifications as the network controller. For example, if the network controller has a PCIe x8 interface then ensure that it is installed into a PCIe slot that can support 8 or more PCIe lanes.

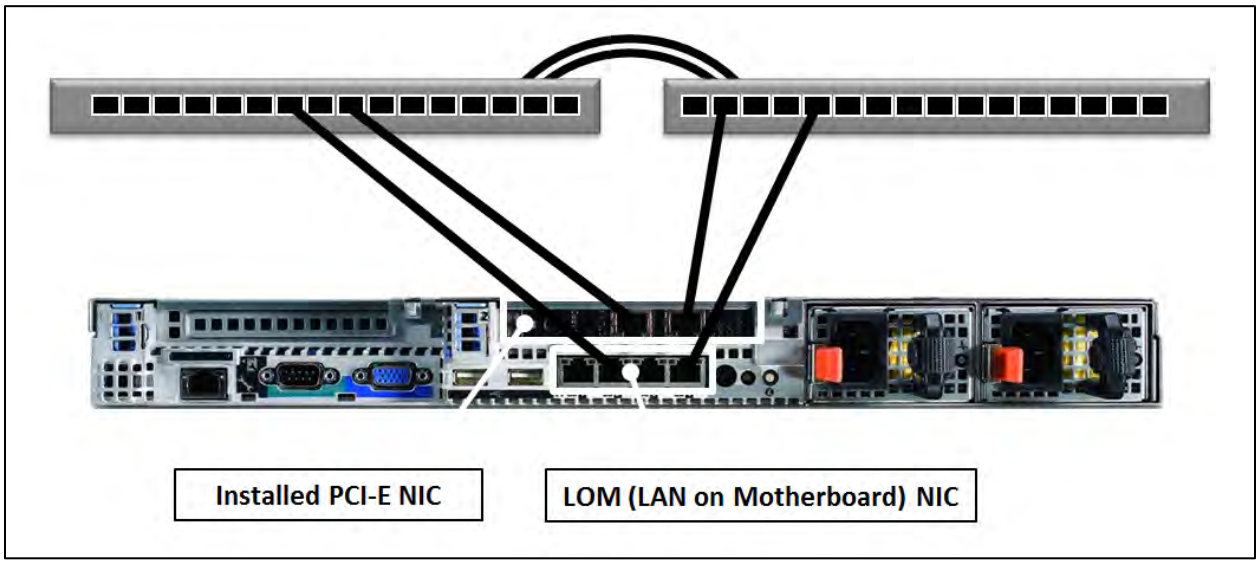


Figure 21 Redundant NIC Connections from Server to SAN using one installed PCI-E NIC and LOM

An R610 server with two additional dual-port PCI-E NICs installed is shown in Figure 22 below. This configuration leaves all four ports on the LOM controller available for other connections.

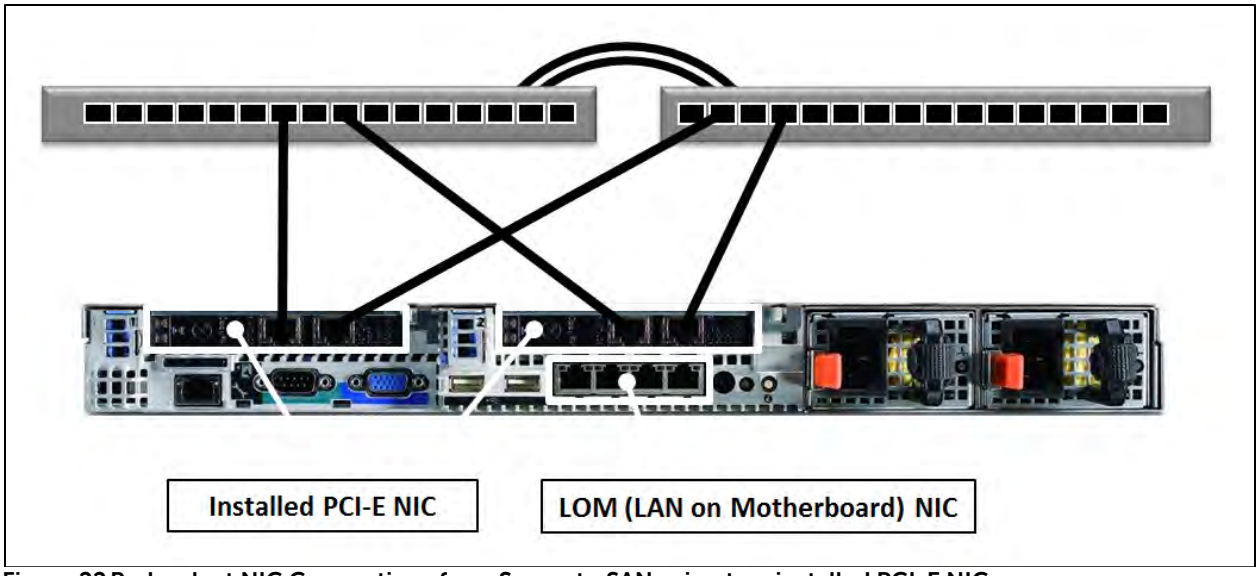


Figure 22 Redundant NIC Connections from Server to SAN using two installed PCI-E NICs

7.3.2 Multi-path I/O

There are generally two types of multi-path access methods for communicating from a host to an external device. For general networking communications, the preferred method of redundant connections is the teaming of multiple NICs into a single, virtual network connection entity. For storage, the preferred method of redundant connection is the use of Multi-Path IO (MPIO). Though some storage solution can and do support either method for iSCSI connectivity, EqualLogic requires the use of MPIO to enable multiple NIC/HBA connections to be utilized for access to an EqualLogic SAN.

EqualLogic MPIO Requirements

The following host port requirements must be met to use MPIO with EqualLogic SANs:

- At least two (2) Ethernet ports are required on each host.
- The host operating system must have a supported MPIO driver or service available.
- The ports used for MPIO cannot be "teamed" to other ports.
- The ports must be the same speed
- The ports must be assigned IP addresses on the same subnet

EqualLogic MPIO General Recommendations

Follow this general set of guidelines for configuring MPIO on a host:

- Configure volume access controls to use standard iSCSI IQN names (See Section 5.2.1 for details). For a more secure configuration you can use the IQN name plus the CHAP security ID.
- On each array enable at least two(2) ports for host connectivity.
- Install the Dell provided MPIO extension features if available for the host operating system.
- For Microsoft Windows, install the Device Specific Module (DSM) found in the Host Integration Toolkit for Windows.
- For VMware vSphere 4.1, install the EqualLogic Multipathing Extension Module.

- For other operating systems, use the native MPIO functionality.

Configuring Microsoft Windows MPIO

Configure Microsoft Windows MPIO with the following initial configuration settings. Customized settings may be required depending on the supported application(s).

- Change the "Subnets included" field to include ONLY the subnet(s) dedicated to the SAN network infrastructure.
- Change the "Subnets excluded" field to include all other subnets.
- The "Load balancing policy" should remain set to the default value of "Least queue depth".
- "Max Sessions per volume slice" should be set to the number of network ports dedicated to SAN subnet (maximum of 4).
- "Max sessions per entire volume" should be set to three (3) times the value of "Max sessions per volume slice" (maximum of 12).
- "Use MPIO for snapshots" should remain at the default setting.
- "Use IPv6 or IPv4" should be set to IPv4 unless your network is configured to use IPv6 as the default communications protocol.

Configuring VMware vSphere 4.1 Multipathing Extension Module (MEM)

Configure the vSphere MEM with the following initial configuration settings. Customized settings may be required depending on the supported application(s).

- Change the "Subnets included" field to include ONLY the subnet(s) dedicated to the SAN network infrastructure.
- "membersessions" should be set to the number of network ports dedicated to SAN subnet (maximum of 4).
- "volumesessions" should be left at the default value of 6.
- "totalsessions" should be left at the default value of 512.

7.3.3 Equallogic iSCI SAN Design

This section will combine all of the SAN components and information provided so far to present a redundant EqualLogic PS Series SAN design. We also include a series of examples illustrating partially redundant and non-redundant SAN configurations.

> **Note:** For FS Series NAS design, the only SAN design patterns presented in this section that we recommend in support of the NAS Service are the fully redundant and fully cabled examples shown in Section 7.3.4. Fully cabled and redundant connection diagrams for the FS7500 are shown in shown in Section 13.1.

The information provided here does not address all of the possible variations in a customer network environment. All information is presented using a set of basic reference designs that make the following assumptions:

- The SAN network is physically isolated from all other network traffic
- The examples are based on best practice design principles.

- Unless otherwise stated, all reference designs will provide end-to-end host to volume redundant paths
- A minimal number of switches will be illustrated to allow for the design concept to be understood. Actual implementations will vary depending on your network infrastructure requirements.
- If sharing physical switches with other non-SAN traffic, we assume all switches are VLAN capable.

7.3.4 Redundant SAN configuration

In a redundant iSCSI SAN, each component of the SAN infrastructure has a redundant connection or path. The following figures show example connection paths necessary to create a Redundant SAN.

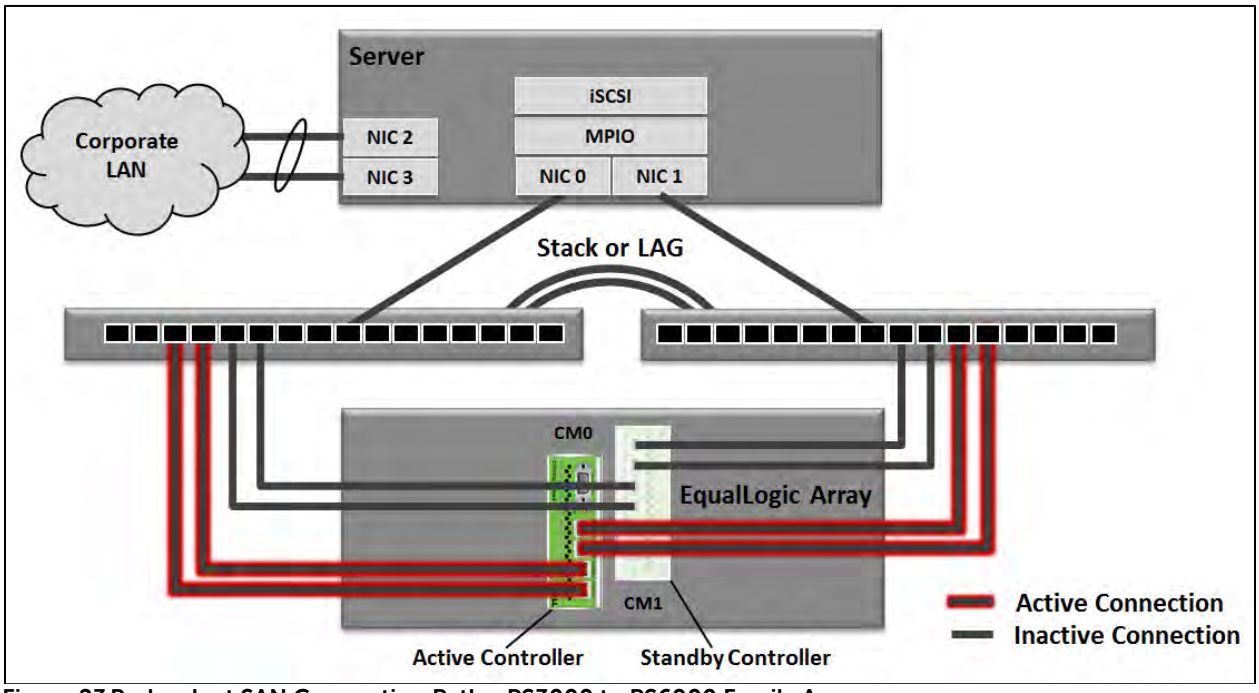


Figure 23 Redundant SAN Connection Paths: PS3000 to PS6000 Family Arrays

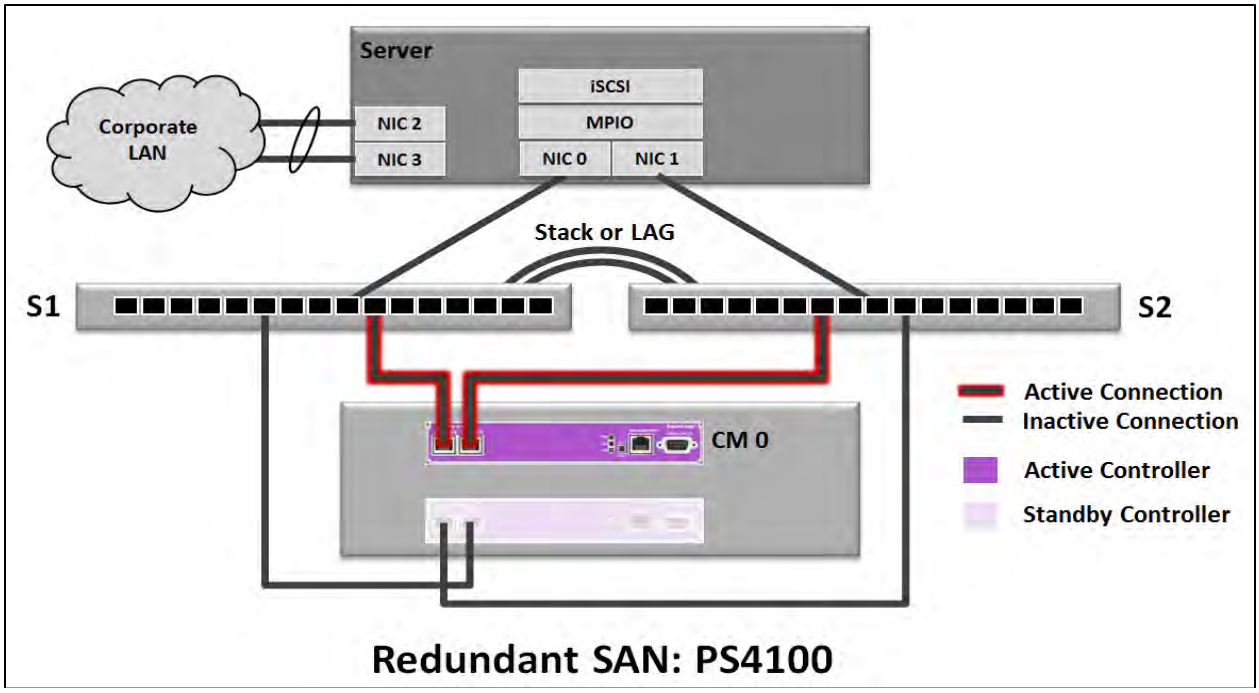


Figure 24 Redundant SAN Connection Paths: PS4100

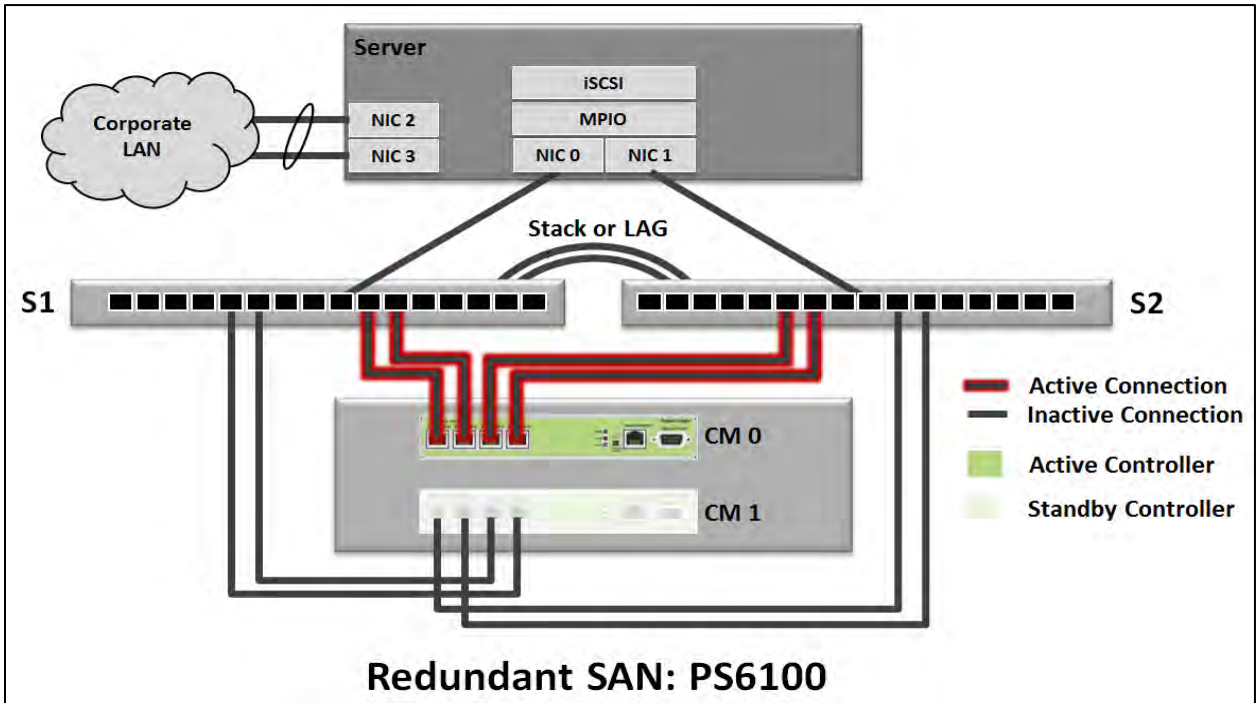


Figure 25 Redundant SAN Connection Paths: PS6100

Note: For a production environment, the configuration examples shown above will protect your access to data. These are the ONLY SAN configurations recommended by Dell.

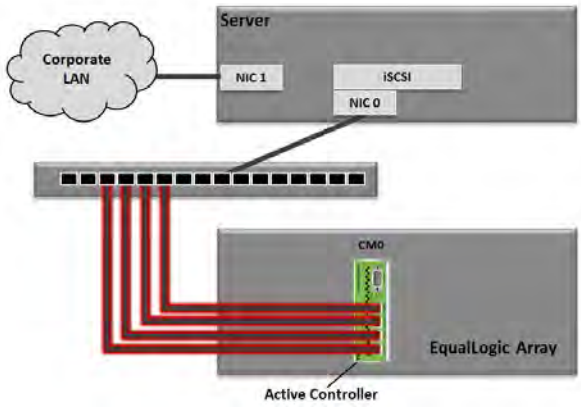
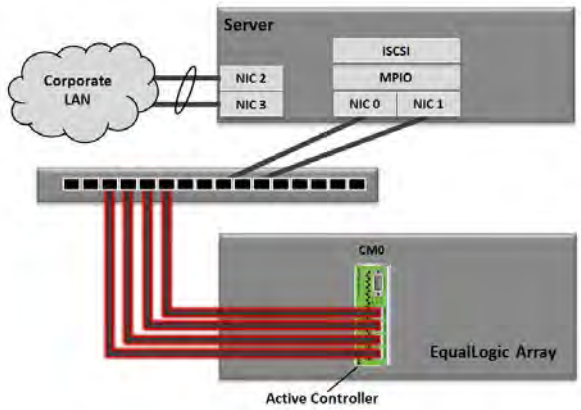
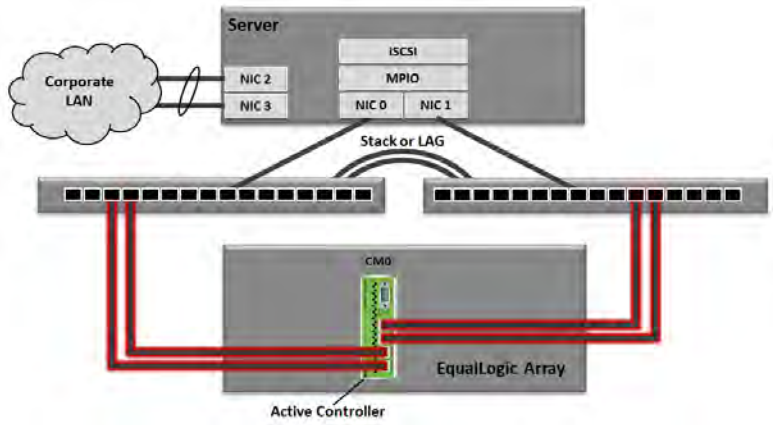
7.3.5 Partially redundant SAN configurations

Each of the SAN configurations shown in this section will allow host connectivity to data stored in the SAN. These configurations are for reference only, and the methods shown apply to both PS3000-PS6000 family controllers and PS4100/PS610 family controllers. **They are not recommended for production deployment since they do not provide end-to-end redundant connection paths.**

7.3.5.1 Single array controller configurations

Table 27 below shows configurations using a single array controller.

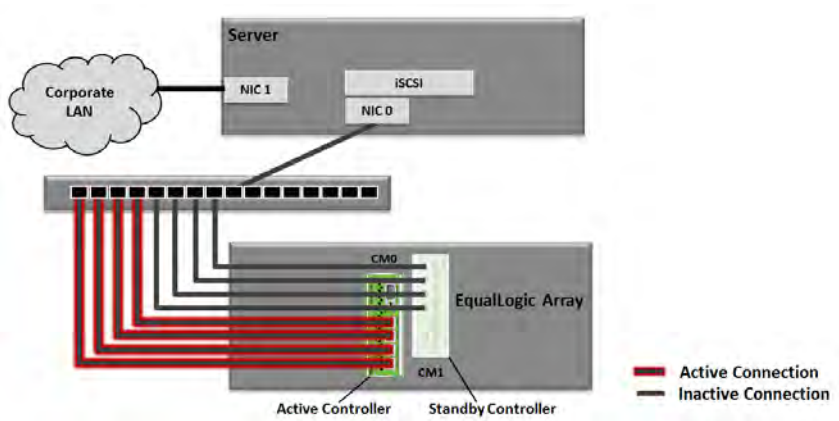
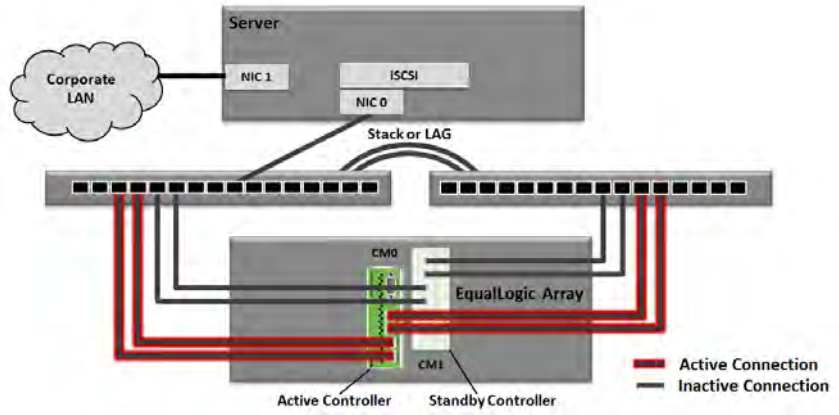
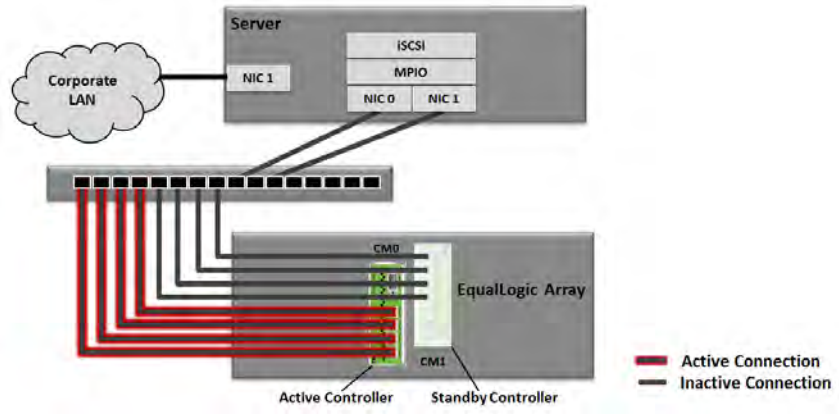
Table 27 Single controller array configurations

Single NIC **Single Switch** **Single Controller**	
Dual NIC **Single Switch** **Single Controller**	
Dual NIC **Dual Switch** **Single Controller**	

7.3.5.2 Dual array controller configurations

You can configure a Dell EqualLogic array to run using dual controllers. Table 28 below shows configurations using a single array controller.

Table 28 Dual controller array configurations

Single NIC **Single Switch** **Dual Controller**	
Single NIC **Dual Switch** **Dual Controller**	
Dual NIC **Single Switch** **Dual Controller**	

7.3.6 Minimum cabling scenarios: PS4100 and PS6100

The vertical port failover feature (described in Section 1.4.4) allows you to cable dual controller PS4100 and PS6100 family arrays for maximum I/O bandwidth and controller redundancy while using

only one half of the available controller ports. The diagrams in Figure 26 and Figure 27 show how to connect these arrays to accomplish this.

> **Note:** The example configurations shown in this section only apply to the PS4100 and PS6100 family arrays and are recommended only when you do not have available SAN switch ports necessary to support fully cabled configurations.

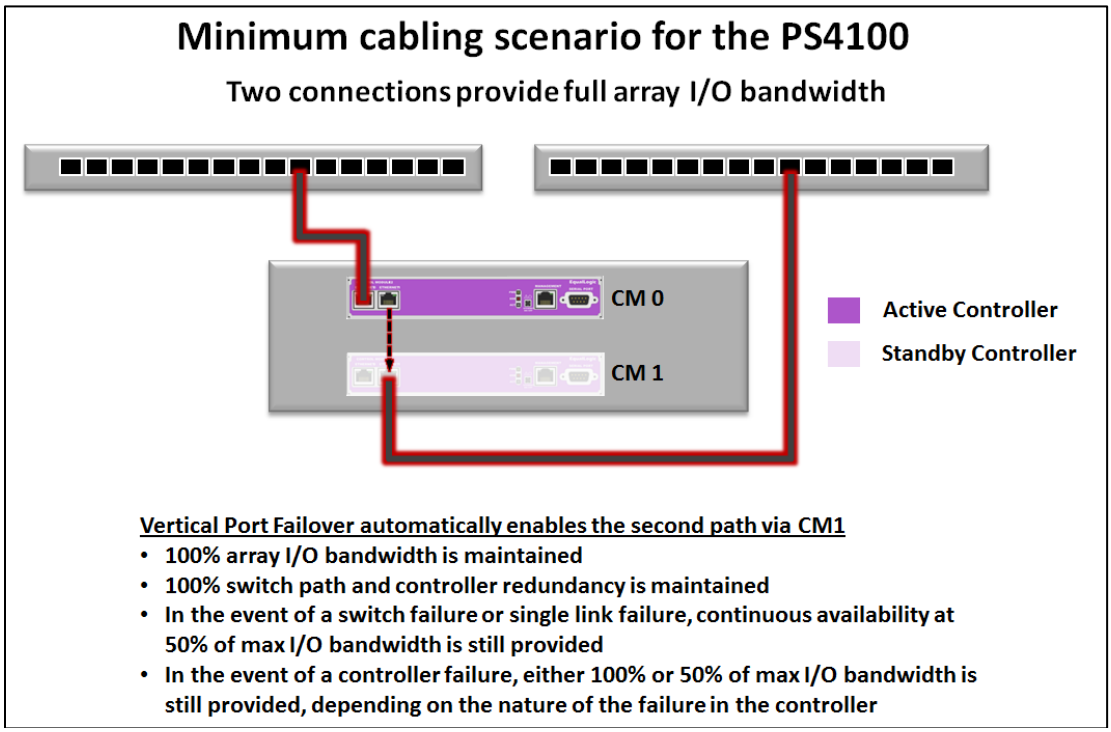

Minimum cabling scenario for the PS4100
Two connections provide full array I/O bandwidth
CM 0
CM 1
Active Controller
Standby Controller
Vertical Port Failover automatically enables the second path via CM1
• 100% array I/O bandwidth is maintained
• 100% switch path and controller redundancy is maintained
• In the event of a switch failure or single link failure, continuous availability at 50% of max I/O bandwidth is still provided
• In the event of a controller failure, either 100% or 50% of max I/O bandwidth is still provided, depending on the nature of the failure in the controller

Figure 26 Minimum cabling scenario: PS4100

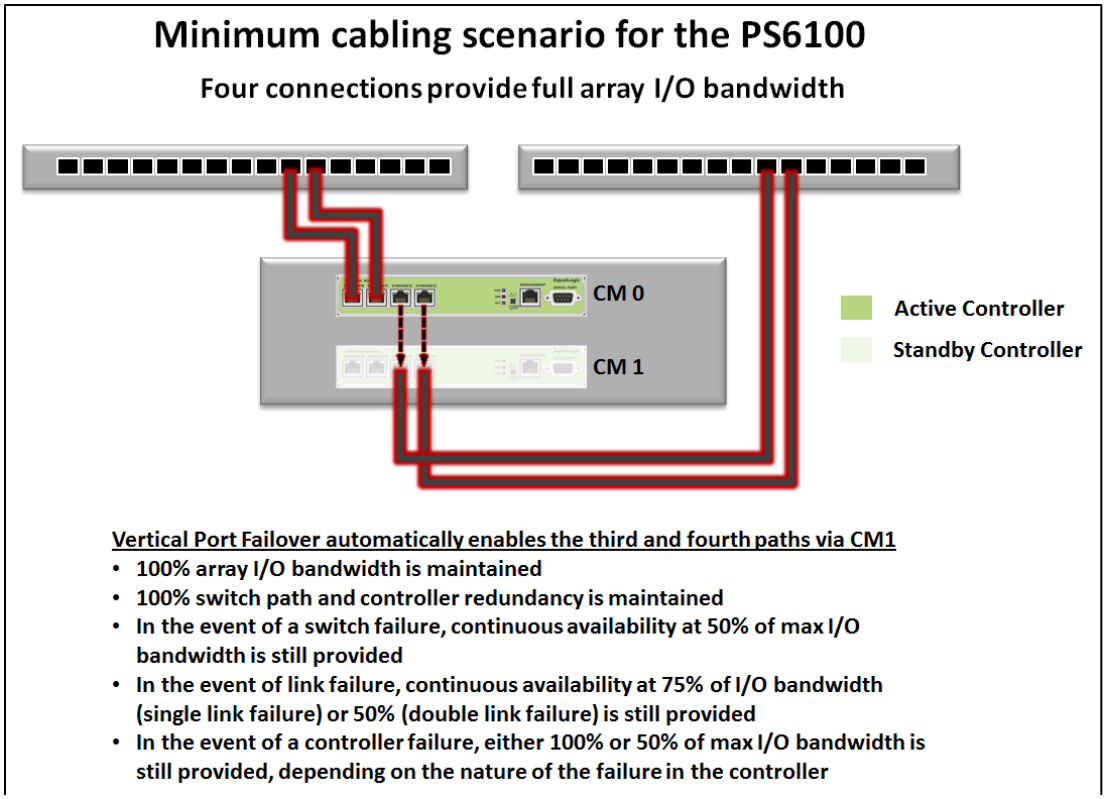

Minimum cabling scenario for the PS6100
Four connections provide full array I/O bandwidth
CM 0
CM 1
Active Controller
Standby Controller
Vertical Port Failover automatically enables the third and fourth paths via CM1
• 100% array I/O bandwidth is maintained
• 100% switch path and controller redundancy is maintained
• In the event of a switch failure, continuous availability at 50% of max I/O bandwidth is still provided
• In the event of link failure, continuous availability at 75% of I/O bandwidth (single link failure) or 50% (double link failure) is still provided
• In the event of a controller failure, either 100% or 50% of max I/O bandwidth is still provided, depending on the nature of the failure in the controller

Figure 27 Minimum cabling scenario: PS6100

8 Mixed speed environments - Integrating 1GbE and 10GbE SANs

With the introduction of 10GbE, there will be situations that require 1Gb arrays and 10Gb arrays coexisting in the same SAN infrastructure. EqualLogic PS Series arrays support operation of 1Gb and 10Gb arrays within the same group. This section summarizes mixed speed SAN design guidelines that are presented in much more detail in the following publications:

- Best Practices for Deploying a Mixed 1 Gb/10 Gb Ethernet SAN using Dell EqualLogic Storage Arrays:
  http://en.community.dell.com/techcenter/storage/w/wiki/2640.deploying-mixed-1-gb-10-gb-ethernet-sans-using-dell-equallogic-storage-arrays-by-sis.aspx
- Integrating EqualLogic PS6x10 Arrays with Existing SANs:
  http://www.equallogic.com/resourcecenter/assetview.aspx?id=9447

The potential advantages in running a mixed speed (1GbE and 10GbE) EqualLogic SAN include:

- Not all of the application workloads on a SAN will require storage I/O performance that the 10Gb arrays provide. Thus, SAN administrators will have additional storage tiering flexibility based on array I/O performance.
- The PS Series Group Manager will allow the SAN administrator to still manage both types of arrays within the same SAN group.
- The ability to mix 1Gb and 10Gb arrays supports seamless operational coexistence during during migration to a 10Gb SAN.

8.1 Design considerations

To properly implement a mixed speed SAN, you must pay close attention to the following design and implementation considerations:

- Ethernet switch feature set, port density and required configuration settings
- Optimization of Rapid Spanning Tree Protocol behavior
- Optimal switch interconnect pattern
- Awareness of I/O workload patterns coming from 1Gb and 10Gb initiators vs. target volume locations in the SAN
- I/O performance implications when using mixed speed vs. segregated speed pools

To create a redundant, mixed speed iSCSI SAN, at a minimum we recommend that you start with dual 1GbE and dual 10GbE switches. Figure 28 shows an example SAN design, where two switches of each type are used.

Referring to Figure 28:

- The design is based on using features provided by the Dell PowerConnect 6248 1Gb Ethernet switch and the Dell PowerConnect 8024 10Gb Ethernet switch.
- The Dell PowerConnect 8024 is not a stackable switch, so a link aggregation group (LAG) is used to create the inter-switch trunk paths.

- Each of the 1Gb switches is configured with one dual-port 10GbE uplink module and one stacking module. The 10GbE uplink modules are used for creating 20Gb LAG uplinks to the 10Gb switches.
- **Split Interconnect** –The 20Gb LAG uplinks between the 1Gb and 10Gb switches are cross-connected so that each 10Gb switch physically connects to both switches in the 1Gb stack.
- **FS7500 Connection Path** – The initial FS Series NAS product (FS7500) is a 1Gb only solution. The NAS Reserve pool for the FS7500 could also be stored in a mixed 1Gb/10Gb PS Series SAN. In this case the FS Series appliance must connect via the 1Gb switch path as shown in Figure 28 and Figure 29.

The 10Gb switches are connected together using multiple 10Gb links. This provides us a path for devices connected directly to the 10Gb switches to communicate without having to go through the slower 1Gb switches.

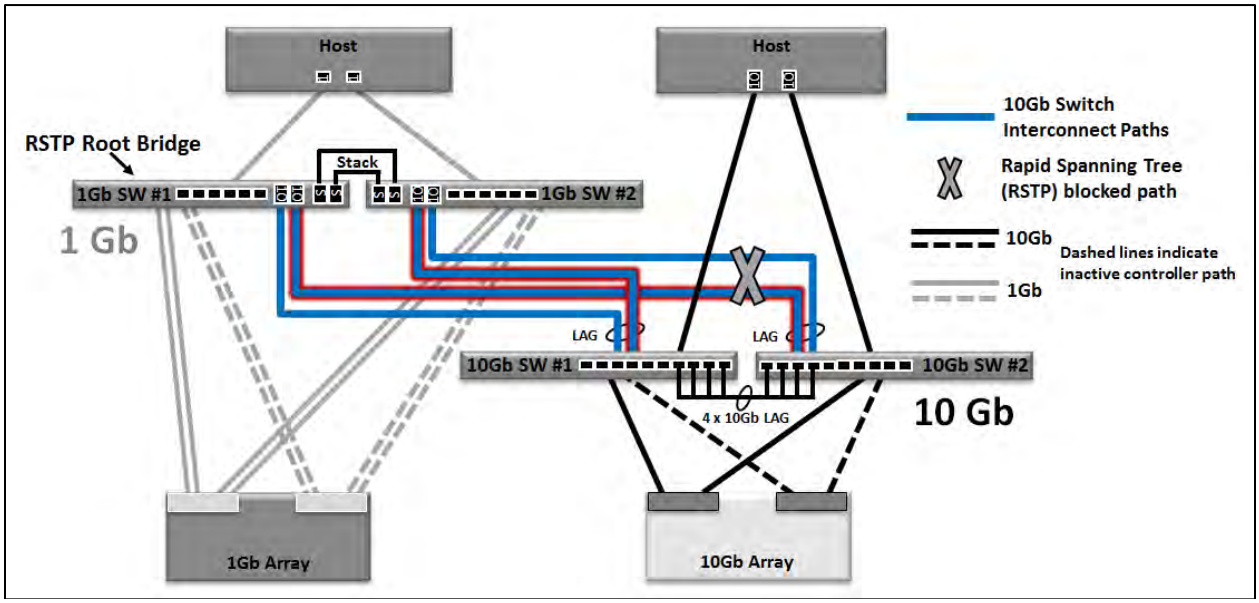


Figure 28 Mixed speed redundant SAN using split interconnect between 1Gb and 10Gb switches

8.1.1 Optimizing Rapid Spanning Tree Protocol behavior

The LAG between the 10Gb switches in Figure 28 creates a loop in the network. Rapid Spanning Tree Protocol (RSTP) will compensate for this by blocking paths as necessary. The optimal spanning tree strategy for this network is to prevent RSTP from blocking the inter-switch trunk between the 10Gb switches, thus causing some 10Gb traffic to traverse the slower 1Gb switches. To accomplish this you need to be aware of which switch is acting as the root bridge in the spanning tree. For the mixed speed SAN design shown in Figure 28 the root bridge is **1Gb SW#1**. Based on this information you can assign a link cost to ensure that the desired link configuration is achieved. Note the location of the RSTP blocked path in Figure 28. We manually assigned a high link cost to one of the 20Gb uplink LAGs so that it became the RSTP blocked path.

The straight uplink pattern must be used when using non-stackable switches

If you are using switches that do not support a stacking mode then you must use the straight interconnect uplink pattern shown in Figure 29. Note the following design differences between the split uplink pattern in Figure 28 and the straight uplink pattern in Figure 29:

- A LAG is used to create the connection between **1Gb SW#1** and **1Gb SW#2**.
- A high rapid spanning tree link cost is assigned to the 1Gb switch LAG (note the location of the RSTP blocked path in Figure 29). Doing this prevents 10Gb inter-switch traffic from having to pass through the 1Gb switch LAG.

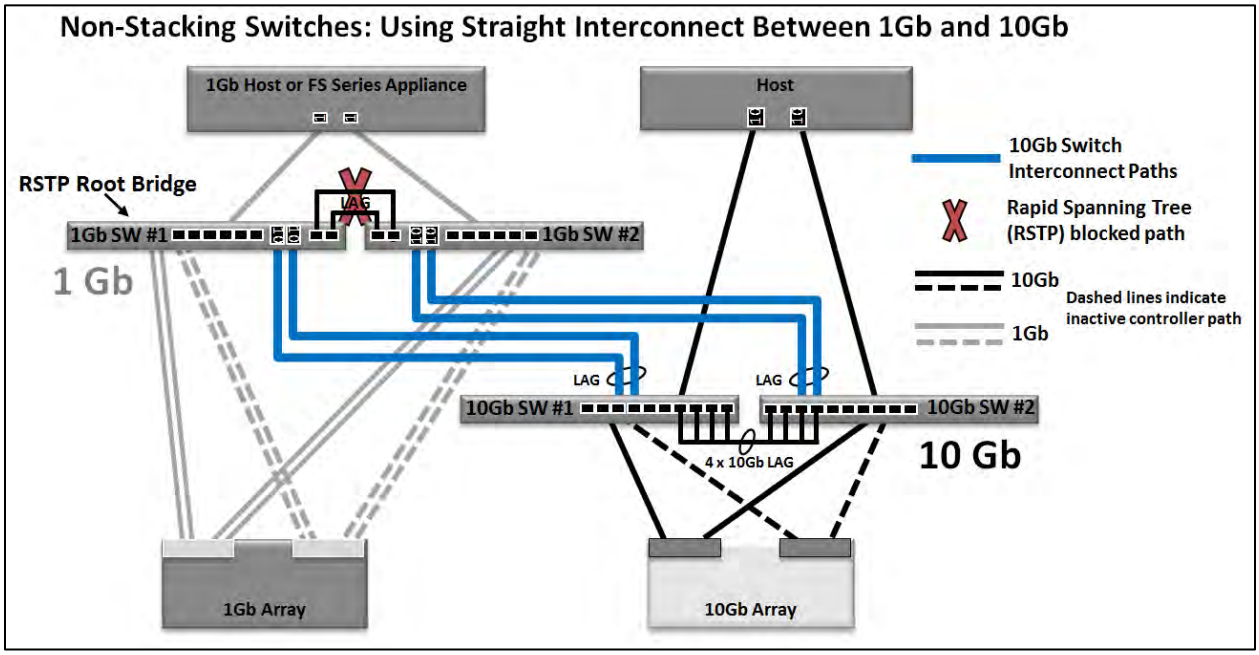


Figure 29 Mixed speed redundant SAN using straight interconnect between 1Gb and 10Gb switches

8.2 Mixed speed SAN best practices

The following list summarizes the important SAN design considerations for integrating 10Gb EqualLogic arrays into existing 1Gb EqualLogic SANs.

- When integrating 10Gb switches into your existing 1Gb environment, how you interconnect the mixed speed switches (split vs. straight uplink) does not have a significant impact on performance as long as the uplinks are sized appropriately to your workloads.
  - If your 1Gb switches are configured as a stack then you should use the split interconnect pattern shown in Figure 28 by default.
  - If your 1Gb switches are not stacked, then you must use the straight interconnect pattern shown in Figure 29.
- When connecting 1Gb switches and 10Gb switches together you must always be aware of where Rapid Spanning Tree is going to block links to make sure that 10Gb traffic (i.e. EqualLogic inter-array data flow) never crosses the 1Gb switch.
- You must configure pools and volumes in a way that minimizes impact to IO performance.

- If you have predominately 1Gb initiators, start upgrading your arrays to 10Gb for comparable or better performance across almost all situations.
- If you have predominately 10Gb initiators, you should only access data and volumes residing on 10Gb arrays (from those initiators). You may see high latency and retransmit rates when 10Gb initiators connect to 1Gb targets.
- When adding 10Gb arrays, place them in separate pools from your 1Gb arrays.

9 Blade server chassis integration

Integrating the PowerEdge M1000e Blade Server Solution (or any third party blade chassis implementation) requires additional SAN design considerations. Each M1000e can support up to three separate networking "fabrics" that interconnect ports on each blade server to a pair of blade I/O modules within each chassis fabric through an intervening chassis midplane interface. Each fabric is associated with different interfaces on a given blade server as described in Table 29. Each blade server has a "LAN on Motherboard" capability that is mapped to the I/O modules located in the Fabric A I/O modules slots on the M1000e chassis and only supports 1Gb or 10Gb Ethernet networking depending on the blade server model. In addition, each blade server has two "mezzanine" sockets for adding additional networking options such as 1Gb or 10Gb Ethernet, Infiniband, or Fibre Channel cards. These mezzanine cards are mapped to either the Fabric B or the Fabric C I/O modules.

Figure 30 illustrates the layout of the three fabric I/O modules located on the back of the M1000e chassis.

Table 29 M1000e fabric mapping

	LOM/MLOM	Mezzanine B	Mezzanine C
Fabric	A	B	C

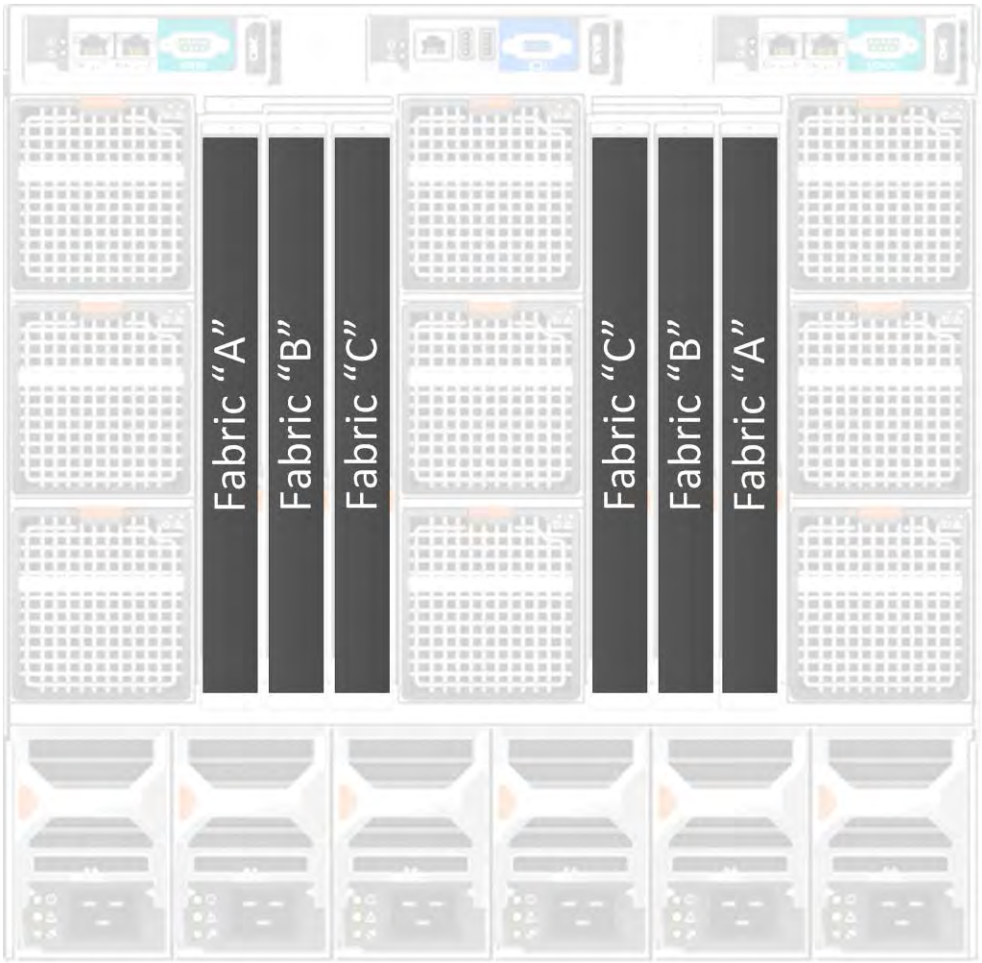


Figure 30 Blade I/O modules and M1000e Chassis

There are three primary methods of integrating blade chassis with EqualLogic PS Series arrays.

- Directly attaching EqualLogic arrays to the blade I/O modules on each chassis.
- Using Pass-Through module to external switches
- Utilizing a two-tier design by creating an external SAN switch infrastructure to host all array connections and using the blade I/O modules as "host access" switches.

For each of these three general SAN design strategies, the user must make design decisions with respect to the type of interconnects to use between the blade I/O modules and/or the external stand-alone switches. Depending on the design recommendation, stacking, link aggregation, or a combination of both types of interconnect may be used. In Section 9.1 we will discuss strategies for directly attaching arrays to blade I/O modules within one or more M1000e chassis. In Section 9.2 we discuss connecting M1000e blade chassis to an external switch using pass-through module. In Section 9.3 we discuss strategies for connecting M1000e blade chassis to an external SAN infrastructure utilizing a two-tier design.

9.1 Designing a SAN using blade chassis I/O modules with arrays directly attached

The following points should be taken into consideration when planning a SAN solution that requires directly attaching EqualLogic arrays to M-Series blade I/O modules:

- Limited number of externally accessible ports typically available on blade I/O modules

  Current M-Series blade I/O modules have limited numbers of externally accessible Ethernet ports. For a 1GbE SAN solutions each array will require up to four ports per I/O module and for a 10GbE SAN solution each array will require up to two ports per I/O module. Based on these maximum requirements, the table below provides a breakdown on the number of externally accessible ports on each model of PowerConnect M-Series blade I/O modules and the maximum number of arrays that could be directly attached to a single M1000e blade chassis.

Table 30 Blade I/O Module options for EqualLogic

	Maximum Available External Ports	Ports Recommended for Interconnect	Arrays Supportable per M1000e
PowerConnect M6220	8x 1GbE	0(Stackable)	2
PowerConnect M6348	16x 1GbE	0(Stackable)	4
PowerConnect M8428-k	8x 10GbE	2x10GbE(PS6010/6510) 3x10GbE(PS4110/6110)	3(PS6010/6510) 5(PS4110/6110)
PowerConnect M8024-k	8x 10GbE	2x10GbE(PS6010/6510) 3x10GbE(PS4110/6110)	3(PS6010/6510) 5(PS4110/6110)
Force10 MXL	2x40GbE Fixed + 2x option slots (2x40GbE, 4xSFP+, 4x10GbaseT)	2x40GbE (Multiple chassis, Stacked)	16(PS4110/6110) 16(14x PS6010/6510) (+ 2x PS-M4110)
Cisco Catalyst Blade Switch 3032	8x 1GbE	4x1GbE	1
Cisco Catalyst Blade Switch 3130G	4x 1GbE	0(Stackable)	1
Cisco Catalyst Blade Switch 3130X	4x 1GbE	0(Stackable)	1
1GbE or 10GbE Pass-Through Module	16	n/a	n/a

- All switches within the SAN infrastructure, including blade I/O modules, must provide a path from all ports hosting array connections to all other ports within the SAN.
- An important networking requirement for EqualLogic SAN infrastructure design is that all switches must be part of the same Layer 2 Ethernet fabric if they will be hosting connections from EqualLogic arrays. For this reason, if arrays are going to be directly attached to one or more blade chassis, then the I/O modules within the Fabric that will be hosting the arrays must be interconnected. Since the I/O modules are not interconnected via the chassis midplane, the only alternative is to use external ports (stacking or Ethernet) to make these inter-switch connections.
- Maximum stack size versus practical stack size
- Depending on the I/O module model, the maximum number of switches allowed in a single stack may be different. In addition, the number of switches supported by the switch may not be optimal for a SAN using EqualLogic arrays. SAN solutions tend to be sensitive to latency and adding multiple hops between a host and target may create situations where latency becomes unacceptable for your application. Testing is advised.
- Each M1000e enclosure will contribute two switches to the stack placing an upward limit on the number of enclosures that can be interconnected into a single SAN network. Regardless of the maximum number of I/O modules supported in a single stack, Dell recommends not going above six switches in the stack due to possible hop related latency.
- M-Series I/O module Stacking Compatibility

- Not all M-Series I/O modules can be stacked together or with external switches. Typically, each M-Series I/O modules model can only stack with modules of the same exact model. It may also be possible to stack M-Series I/O modules with some "stack compatible" stand-alone PowerConnect switch models. The table below provides stack compatibility information.

Table 31 Stack Compatibility for M-Series I/O modules

	Stack Compatible Switches	Maximum Stack Size	Recommended Stack Size
PowerConnect M6220	PowerConnect M6220 PowerConnect 6224 PowerConnect 6248	10	4
PowerConnect M6348	PowerConnect M6348 PowerConnect 7024 PowerConnect 7048	12	6
PowerConnect M8428-k	N/A	N/A	N/A
PowerConnect M8024-k	M8024-k	6	6
Force10 MXL	Force10 MXL	6	4
Cisco Catalyst Blade Switch 3032	N/A	N/A	N/A
Cisco Catalyst Blade Switch 3130G	Catalyst BS3130G Catalyst BS3130X	8	4
Cisco Catalyst Blade Switch 3130X	Catalyst BS3130G Catalyst BS3130X	8	4

SAN Design for a Single M1000e Enclosure

For a single M1000e enclosure, there is a requirement to use two blade I/O modules for redundancy. Depending on the blade I/O module model, you may use either stacking or link aggregation as described earlier in Table 30. Where stacking is supported, this will most likely be the primary method of interconnecting the blade I/O modules to create a single switch fabric. SAN switches can be located in any of the three enclosure fabrics and it is not required that the blade I/O modules be in the same enclosure fabric, though it is typical that the two I/O modules would be in the same enclosure fabric.

Placing the blade I/O modules in the same fabric does remove one aspect of high availability in that each blade server will have both of the SAN ports located on the same fabric mezzanine card. This creates a potential single point of failure if the mezzanine card as a whole fails.

One alternative configuration would be to place the two blade I/O modules into enclosure slots associated with two different enclosure fabrics (B1 and C1 for example). This has the advantage that each blade server will have its Ethernet ports connected to these blade I/O modules on two different mezzanine cards. This ensures that even if a single mezzanine card fails, there is still an active port on the SAN network.

Table 32 and Table 33 illustrate the stacking strategies between the two I/O modules in each of these two configurations as described.

Table 32 Single M1000e enclosure stacking single fabric

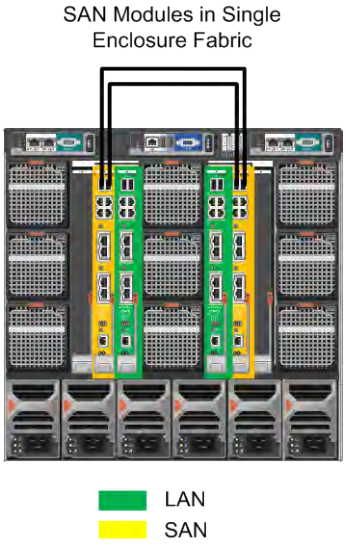
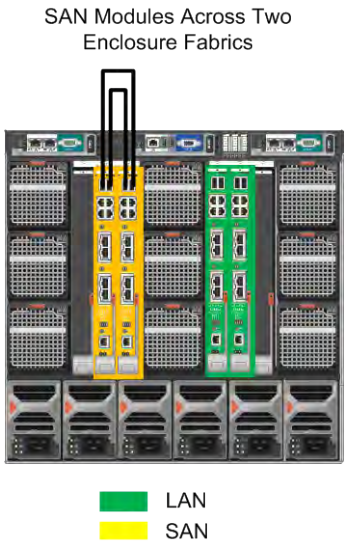
Single Fabric Stacked Configuration	
	
Advantages	**Concerns**
• Consistent M1000e fabric management that adheres to M1000e recommended practices for blade IO configuration • Reduces administration overhead	• All blade server Ethernet ports used for SAN reside on the same mezzanine card resulting in a potential single point of failure • Upgrading switch FW will require scheduled downtime for SAN network

Table 33 Single M1000e enclosure stacking dual fabric

Dual Fabric Stacked Configuration	
	
Advantages	**Concerns**
• Ensures that Ethernet SAN ports on each blade server will be distributed across two different mezzanine cards for a more highly available solution	• Does not adhere to recommended practices for M1000e blade IO configuration • Upgrading switch FW will require scheduled downtime for SAN network

As discussed earlier, one of the concerns when creating a SAN that consists of a single stack is that a single stack does not provide for a highly available solution when attempting to perform switch maintenance such as switch firmware updates. If there is a requirement that the SAN solution cannot be brought down for regularly scheduled maintenance, then stacking may not be the right option and you should consider link aggregation methods for switch interconnect.

Table 34 and Table 35 illustrate the non-stacking strategies between the two I/O modules in each of these two configurations as previously described.

Table 34 Single enclosure link aggregation single fabric

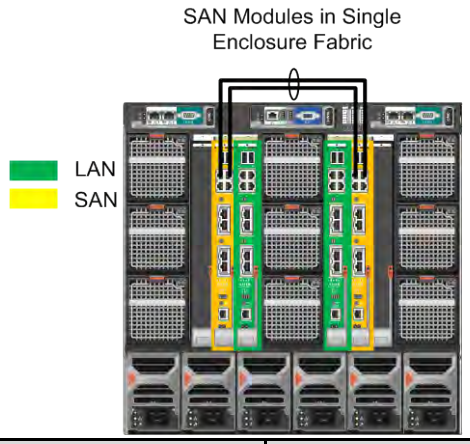
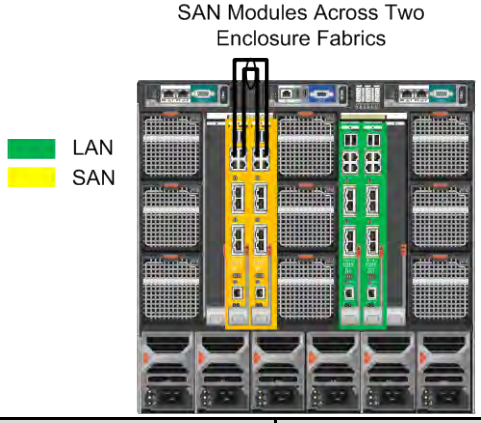
Single fabric non-stacked configuration	
	
Advantages	**Concerns**
• Consistent M1000e fabric management that adheres to M1000e recommended practices for blade IO configuration • Switch FW can be upgraded without requiring network to be brought offline	• All blade server Ethernet ports used for SAN reside on the same mezzanine card resulting in a potential single point of failure • Spanning Tree must be considered if uplinking SAN to external switches or additional M1000e enclosures

Table 35 Single Enclosure Link Aggregation Dual Fabric

Dual fabric non-stacked configuration	
	
Advantages	**Concerns**
• Ensures that Ethernet SAN ports on each blade server will be distributed across two different mezzanine cards for a more highly available solution • Switch FW can be upgraded without requiring network to be brought offline	• Does not adhere to recommended practices for M1000e blade IO configuration • Spanning Tree must be considered if uplinking SAN to external switches or additional M1000e enclosures

The 10GbE PowerConnect M-Series I/O modules do not provide a dedicated stacking interface and must be interconnected using available, external Ethernet ports in conjunction with a link aggregation protocol such as LACP or "front-port" stacking. Due to the limited number of external ports available on Dell's PowerConnect M-Series blade I/O modules, SAN growth can be limited. For 10Gb SAN solutions that will require multiple arrays and/or multiple M1000e chassis, it is recommended that you do not consider directly attached arrays and follow the guidelines described in Section 9.2.

If reducing administrative overhead is the goal, the Direct Attached design with a inter-fabric stack is the best option, with the storage direct attached design requiring fewer cables thatn the host pass-through design.

Direct Attached – Single Chassis Stacked and/or LAG

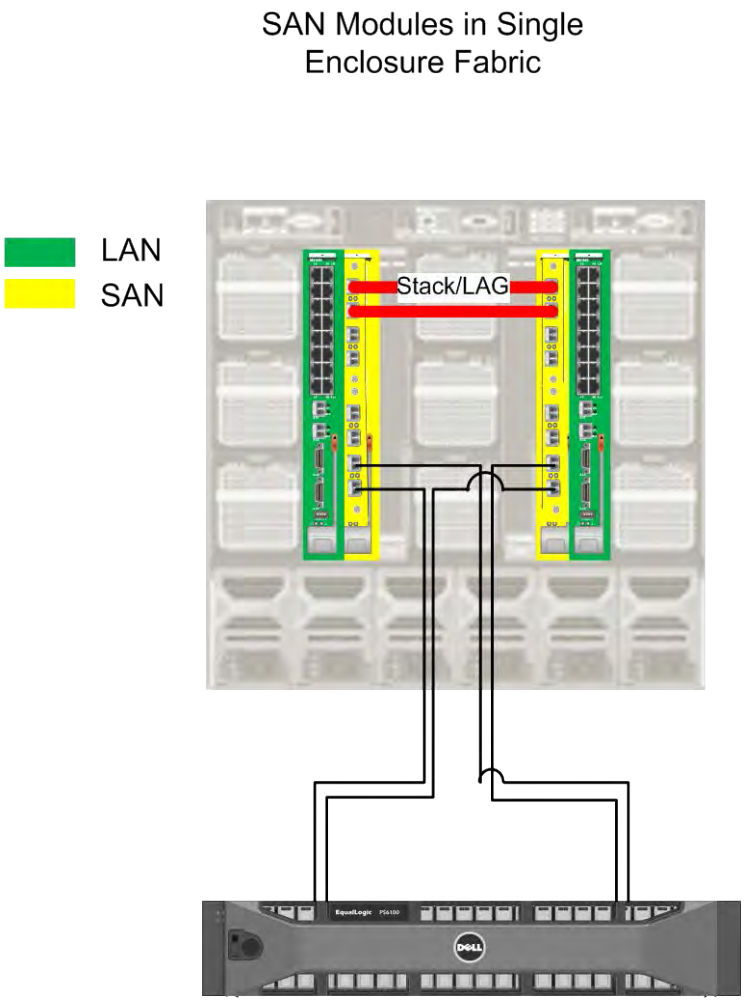


Figure 31 SAN modules in single enclosure fabric

9.1.1 SAN design for multiple M1000e enclosure

For a multiple M1000e enclosure SAN solutions, there are two primary strategies to consider, the first – described here – is directly attaching arrays to the individual M1000e chassis and interconnecting the chassis M-Series I/O modules in each chassis, and the second strategy – described in Section 9.2 – is to build an external network for array consolidation and using chassis M-Series I/O modules as

"host access" switches. In most cases, the latter method of SAN design is preferable to the former, but this section will describe strategies to build a SAN that consists of multiple M1000e enclosures.

As in the previous section, there are two methods that must be considered when interconnecting M-Series I/O modules: stacking and non-stacking. Due to the limited number of external ports available on M-Series I/O modules, stacking should be the primary method of interconnecting M1000e enclosures as these external ports will need to be used for attaching EqualLogic arrays.

As just indicated stacking of M-Series I/O modules should be the primary method of building a multiple enclosure SAN with arrays directly attached to the M-Series I/O modules. As with the single enclosure discussion, stacking has some advantages and disadvantages from a network management perspective and network availability perspective. Take this into consideration before deciding to implement a multi-enclosure SAN using directly attached arrays.

Table 36 illustrates the primary method for building a multi-enclosure SAN with directly attached arrays. Each M1000e enclosure should be able to host one to four arrays depending on the M-Series I/O module used. By stacking each enclosure's fabric I/O modules into a single stack, a single SAN network infrastructure can be constructed.

Table 36 Single stack multiple M1000e enclosure direct connect SAN

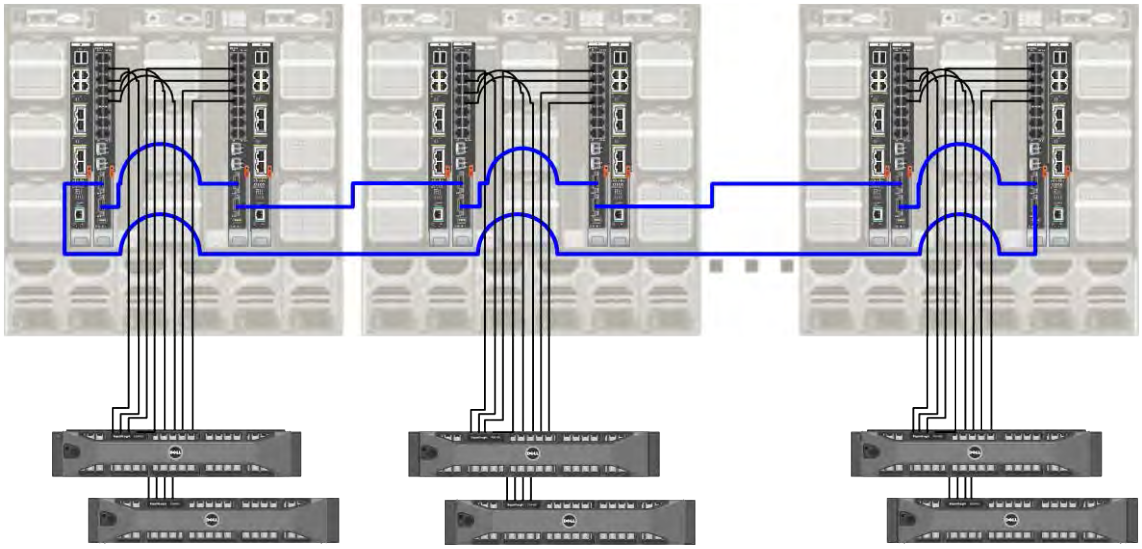
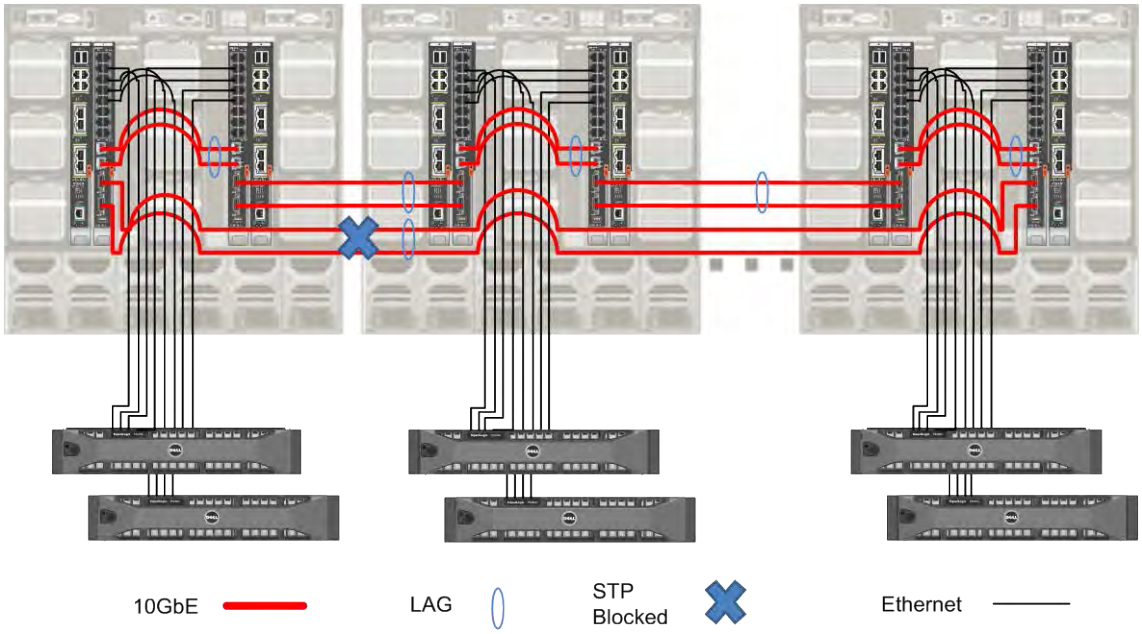
Single stacked, multiple M1000e enclosure configuration	
	
Advantages	**Concerns**
• Simplified switch management • High-bandwidth interconnect	• Not highly available, single stack solution cannot be maintained without scheduled downtime. • Additional scalability will require SAN network redesign and possible downtime.
Other notes	
• Limited scalability due to potentially excessive hop-counts and latency. Recommend no more than 2-3 enclosures in this design	

Table 37 Non-stacked, multiple M1000e enclosure direct connect SAN

Non-Stacked, Multiple M1000e Enclosure Configuration	
	
Advantages	**Concerns**
• Required for non-stacking M-Series I/O modules to support direct attached arrays. • Array firmware can be updated without SAN being brought down for maintenance	• Spanning tree protocol will block one link to prevent a loop • Spanning tree may change if M-series I/O modules are uplinked to other switches (such as when attempting replication to remote data center). • Requires more cables than stacked solutions
Other notes	
• Should only use 10GbE links for this strategy to ensure adequate bandwidth between modules and enclosures. • Limited scalability due to potentially excessive hop-counts and latency. Recommend no more than 2-3 enclosures in this design	

If large SAN solutions are required, the design just discussed may not be the adequate. While the stacking interfaces on the M-series I/O modules is higher than we could obtain using a link aggregation group consisting of 1Gb ports, it is still somewhat limited. Also, if you foresee needing more M1000e enclosures than the two to three enclosure maximum recommended, then a different SAN design strategy should be considered.

9.2 Designing a SAN using blade Pass-through module

This SAN design include configurations where the blade server host ports are directly connected to TOR switches using 1GbE pass-through IOM in the M1000e blade chassis. The storage ports are also connected to the TOR switches,

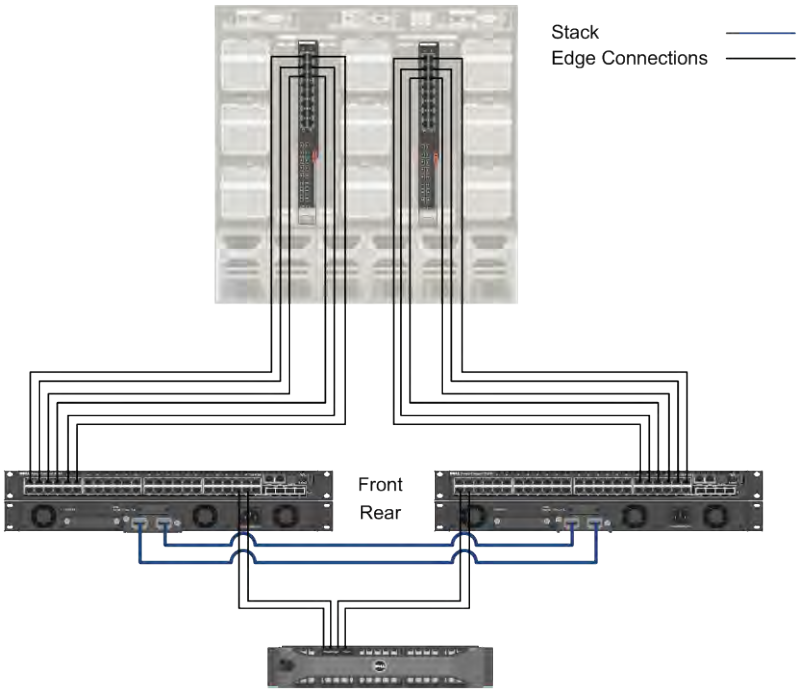
Note that because an ISL stack is not a recommended configuration due to a lack of SAN availability during stack reloads, only one single switch tier design with ISL stack was tested – Blade IOM switch only – and the TOR switch only with ISL stack design was excluded.

This SAN design provides sufficient ISL bandwidth between the two switches by creating a LAG between the external switches. Since there is only a single tier of switches, there is no uplink from the blade IOM switches. The remaining ports on each switch can accommodate the connection of multiple PS Series array members. The host/storage port ratio with the maximum number of array members is 1:1.

The following diagram illustrates how to directly connect the two switches and how the two switches are connected by an ISL LAG. Also note that the corresponding port on the passive controller is should be connected to a different switch than the port on the active controller, ensuring that the port-based failover of the array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

Pass-Through modules are supported for use with EqualLogic SAN solutions. The Pass-Through module provides a simple, direct path from each blade server's optional Ethernet mezzanine card to an externally accessible port. These ports can then be connected to one or more external switches that are hosting PS Series arrays.

Table 38 M-Series I/O module Pass-Through to external switch infrastructure

M-Series I/O module Pass-Through to external switch infrastructure	
 * Single array, Primary array connections only are shown. Additional arrays and standby controller ports connections should be made in a similar fashion following cabling guidelines in Section 2.3	
Advantages	**Concerns**
• Fewer switch hops • Fewer switches to manage • SAN can support M1000e enclosures and rack or tower stand-alone servers • SAN traffic isolated outside of M1000e enclosure • Less expensive that multi-tiered switch solution • Supports non-dell branded switching as a single vendor solution • No firmware to update on the pass-through modules.	• Complex cabling in a rack environment.
Other Notes	
• Additional M1000e enclosures are added in similar fashion to diagram above • Supported with stacked or non-stacked external SAN infrastructures • This strategy works for a single fabric or dual fabric M-Series I/O module configuration (See Section 0 for details) • This is the simplest choice if external switches are from a different vendor.	

9.3 Designing a SAN using blade chassis I/O modules as host access to external switches for array connection

When attempting to have multiple M1000e blade chassis connected to EqualLogic SAN arrays, or if there is a need to also attach traditional rack or tower servers to the SAN, it is strongly recommended that you consider creating a SAN infrastructure that does not directly attach arrays to the blade I/O modules, but rather consider using a set of external switches to host array connections.

The basic strategy is a slight modification from the basic SAN design guidelines used for SANs that do not have M1000e enclosures discussed in Section 7.3.3 Equallogic iSCI SAN Design, so this discussion will start with that basic SAN reference architecture. The difference will be the connection of each M1000e enclosure's M-Series I/O modules to at least two switches in the external SAN infrastructure using either Link Aggregation (the primary method) or Stacking interfaces – if the external switches are stack compatible (See Stack Compatibility for M-Series I/O modules). Even if the M-Series I/O modules and the external switches are stack compatible, it is recommended that stacking not be used due to the inability to upgrade switch firmware without taking the SAN network offline.

The following tables provide various methods for uplinking M1000e enclosures to an external SAN network infrastructure and provide insights into the advantages and concerns that may affect your configuration, depending on your business environment. In the end, it depends on your specific needs to determine which configuration is best for your solution.

Table 39 Horizontal Stack with Vertical LAG

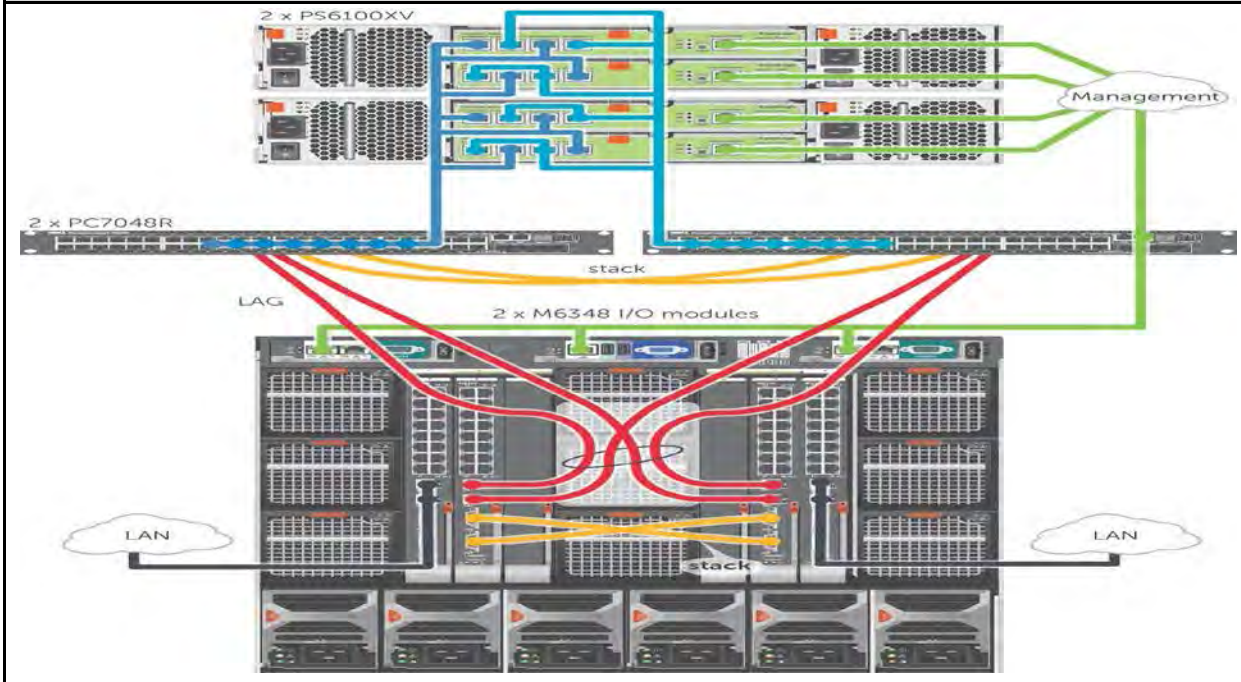
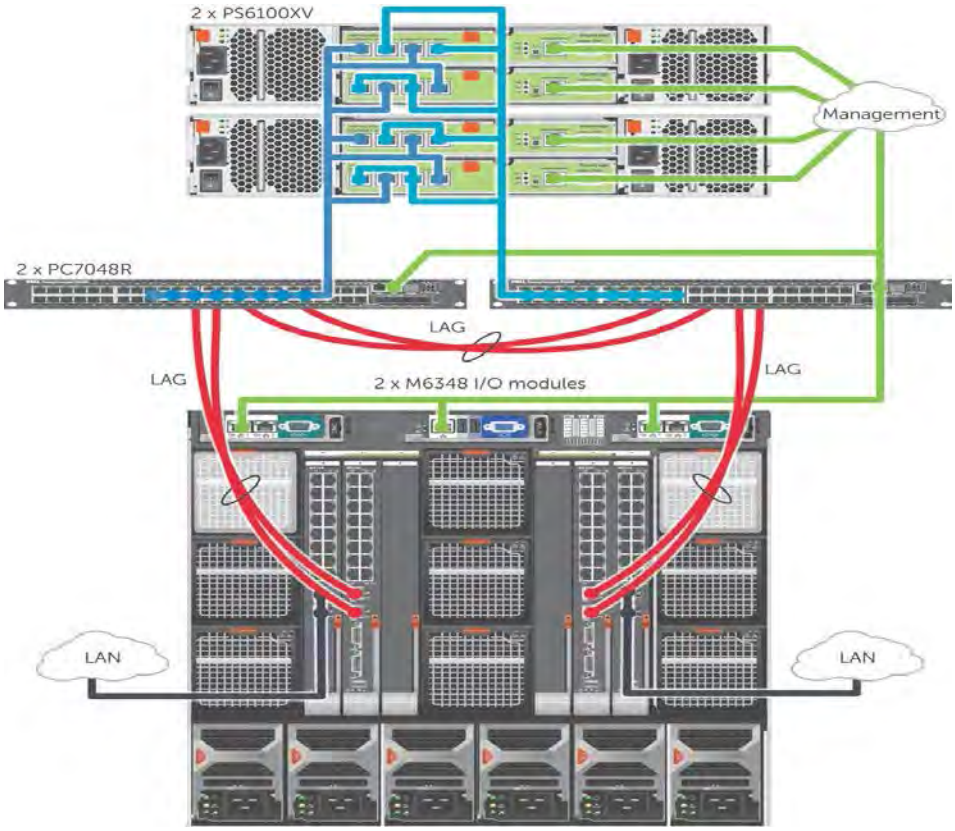
Connecting M1000e to Stacked external switch infrastructure	
	
Advantages	**Concerns**
• Support configuring LAGs between M-series I/O modules across both external switches in stack (cross-switch LAG) • SAN Array-to-array traffic is isolated outside of the M1000e enclosures • Supports ability to scale arrays • Supports ability to scale host. • Supports sharing SAN with non-M1000e servers. • External stacked switches can be expanded by adding additional switches to stack as more ports are needed. • Can support different vendor switches in the external tier	• Must bring external switch stack down to apply switch firmware updates. • Must bring internal switch stack down to apply switch firmware updates.
Other notes	
• If available, use 10Gb links between M-series I/O modules and external switches • Additional M1000e enclosures are added in similar fashion to diagram above.	

Table 40 3-Way LAG

Connecting M1000e to Non-Stacked external switch infrastructure


Advantages	Concerns
• Switch Firmware updates can be made on all switches without downtime • SAN Array-to-array traffic is isolated outside of the M1000e enclosures • Supports ability to scale arrays • Supports ability to scale host. • Supports sharing SAN with non-M1000e servers. • Can support different vendor switches in the external tier	• Limited scalability of LAG bandwidth. • Spanning Tree may block some links if a loop is created. • All switches must be managed separately

Other Notes
• Recommend using 10GbE or larger links within inter-switch LAGs to provide better bandwidth scalability. • Care needs to be taken to ensure that addition of switches does not cause a spanning tree reconfiguration. • Do not interconnect the M-series I/O modules together. • Additional M1000e enclosures are added in similar fashion to diagram above.

Table 41 4-Way Stack

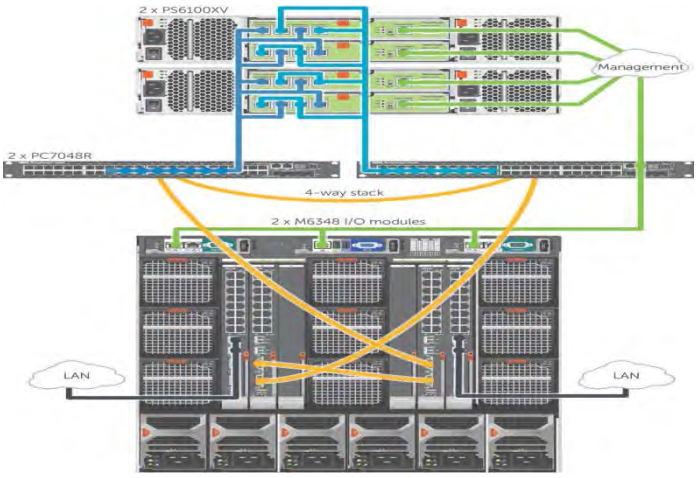
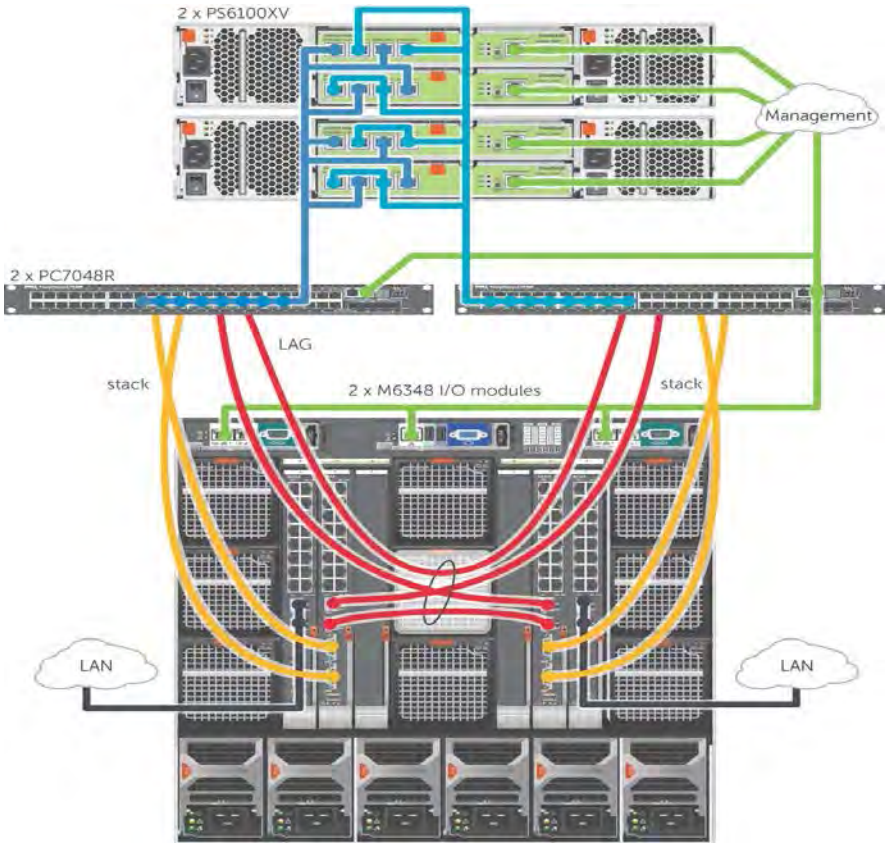
Connecting M1000e to Non-Stacked external switch infrastructure	
	
Advantages	**Concerns**
• SAN Array-to-array traffic is isolated outside of the M1000e enclosures • Fixed hops between hosts and array volumes regardless of number of M1000e enclosures • Supports ability to scale arrays • Supports ability to scale host. • Supports sharing SAN with non-M1000e servers. • Ease of Administration	• Must bring external switch stack down to apply switch firmware updates. • Must bring internal switch stack down to apply switch firmware updates. • Switches must be stack compatible
Other Notes	
• Recommend using 10GbE or larger links within inter-switch LAGs to provide better bandwidth scalability. • Additional M1000e enclosures are added in similar fashion to diagram above.	

Table 42 Vertical Stack with Horizontal LAG

Connecting M1000e to Non-Stacked external switch infrastructure


Advantages	Concerns
• Switch Firmware updates can be made on all switches without downtime • Fixed hops between hosts and array volumes regardless of number of M1000e enclosures • Supports ability to scale arrays • Supports ability to scale host. • Supports sharing SAN with non-M1000e servers. • High Availability	• Limited scalability of LAG bandwidth. • Spanning Tree may block some links if a loop is created.

Other Notes
• Recommend using 10GbE or larger links within inter-switch LAGs to provide better bandwidth scalability. • Care needs to be taken to ensure that addition of switches does not cause a spanning tree reconfiguration. • Do not interconnect the M-series I/O modules together. • Additional M1000e enclosures are added in similar fashion to diagram above.

Table 43 Dual M-Series Stacks to Stacked external switch infrastructure

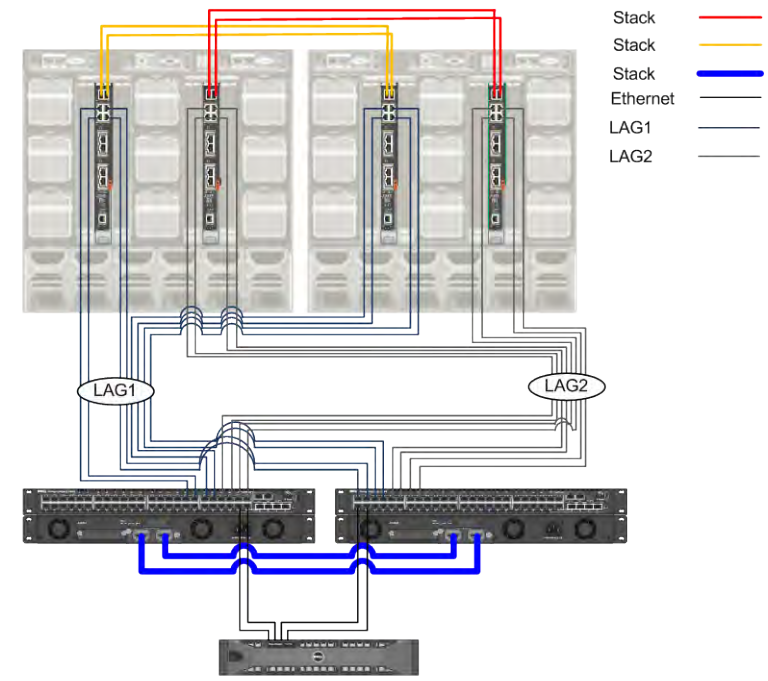
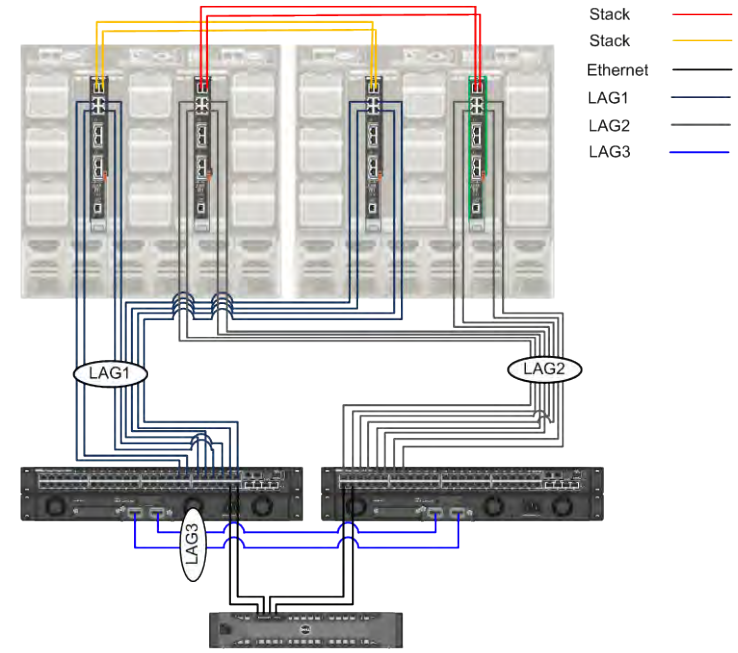
M-Series I/O module Stack(s) to Stacked external switch infrastructure	
 * Primary array connections only are shown. Standby controller ports should be connected in a similar fashion following cabling guidelines in Section 2.3	
Advantages	**Concerns**
• Each M-Series I/O module stack can be updated independently without bringing SAN down. • SAN Array-to-array traffic is isolated outside of the M1000e enclosures • Fixed hops between hosts and array volumes regardless of number of M1000e enclosures • Supports ability to scale arrays • Supports ability to scale host. • Supports sharing SAN with non-M1000e servers. • Fewer "switches" to manage since each stack is managed as a single switch	• Firmware updates on external stack requires scheduled downtime • Limited scalability of LAG bandwidth. • Spanning Tree may block some links if a loop is created.
Other Notes	
• If available, use 10Gb links between M-series I/O modules and external switches • M-series I/O module stacks do not need to be interconnected since no arrays are attached directly to M-series I/O modules. • Additional M1000e enclosures are added in similar fashion to diagram above. Enclosure scalability is greater if 10Gb links are used between M-series I/O module stacks and the external stack • This strategy works for a single fabric or dual fabric M-Series I/O module configuration (See Section 0 for details)	

Table 44 M-Series I/O module Stack(s) to Non-Stacked external switch infrastructure

M-Series I/O module Stack(s) to Non-Stacked external switch infrastructure	
 * Primary array connections only are shown. Standby controller ports should be connected in a similar fashion following cabling guidelines in Section 2.3	
Advantages	**Concerns**
• Each M-Series I/O module stack can be updated independently without bringing SAN down. • SAN Array-to-array traffic is isolated outside of the M1000e enclosures • Fixed hops between hosts and array volumes regardless of number of M1000e enclosures • Supports ability to scale arrays • Supports ability to scale host. • Supports sharing SAN with non-M1000e servers. • Fewer "switches" to manage since each stack is managed as a single switch	• Firmware updates on external stack requires scheduled downtime • Limited scalability of LAG bandwidth. • Spanning Tree may block some links if a loop is created.
Other Notes	
• If available, use 10Gb links between M-series I/O modules and external switches • M-series I/O module stacks do not need to be interconnected since no arrays are attached directly to M-series I/O modules. • Additional M1000e enclosures are added in similar fashion to diagram above. Enclosure scalability is greater if 10Gb links are used between M-series I/O module stacks and the external stack • This strategy works for a single fabric or dual fabric M-Series I/O module configuration (See Section 9.1 for details)	

10 Fluid File system

The Dell EqualLogic FS Series NAS Appliance adds scale-out unified file and block Network Attached Storage (NAS) capabilities to any[2] EqualLogic PS Series iSCSI SANs. The key design features and benefits provided by the EqualLogic FS Series Appliance include:

- A scalable unified (block and file), virtualized IP based storage platform.
- Seamless capacity and performance scaling to meet growing storage demands. By increasing the number of EqualLogic arrays and FS Series controllers you can scale both storage capacity and I/O performance as needed.
- A single configuration and management UI (the EqualLogic Group Manager) from which both block and file storage features are controlled.
- The Dell Fluid File System (FluidFS) is a high-performance scale-out file system capable of presenting a single file-system namespace through a virtual IP address, regardless of the number of NAS appliances in the cluster. FluidFS is designed to optimize file access performance and hardware utilization while eliminating capacity constraints.

Additionally, the FS Series Firmware V2.0 adds the following features:

- **Support for FS7600 and FS7610 NAS appliances**
- **NAS container replicaton:** Version 2.0 of the FS Series Firmware supports NAS container replication, in which a file-system-consistent copy of a container is replicated from one NAS cluster to another. NAS container replication requires FS Series Firmware Version 2.0 on the EqualLogic FS Series NAS appliances.
- **NAS antivirus service for CIFS shares:** NAS clusters support virus scanning by using an antivirus server. You can configure up to four antivirus servers to scan files stored in CIFS shares (NFS is not supported). The NAS cluster solution contains integration with industry standard ICAP-enabled antivirus software to ensure files written from CIFS clients are virus-free. The antivirus host must run Symantec ScanEngine 5.2, which is ICAP-enabled.
- **New monitoring features:** Version 2.0 of the of the FS Series Firmware introduces support for the following new monitoring panels:

    - NAS Snapshot Schedules – Provides information about NAS container snapshot operation according to the regular schedules that you configure by using the Create Schedule (snapshot) activity. You can enable, disable, modify, and delete schedules from this panel.
    - NAS Replication Schedules – Provides information about NAS container replication operations according to the regular schedules that you configure by using the Create Schedule (replica) activity. You can enable, disable, modify, and delete schedules from this panel.
    - Outbound NAS Replication – Provides information about replicated NAS containers where the source container exists on the local group, replicating to a replica container on a partner group.
    - Inbound NAS Replication – Provides information about replicated NAS containers where the replica container exists on the local group, replicated from a source container on a partner group.

[2] Any new or existing EqualLogic PS Series array running controller firmware version 5.1 or later

10.1 FS Series architecture

The FS Series appliance connects to an EqualLogic PS Series SAN via standard iSCSI connection paths with iSCSI based block level I/O taking place between the FS Series appliance initiators and the target volumes on the PS Series arrays. The FS Series controllers host the FluidFS based NAS file system and all front-end client protocol connection management and load balancing, as well as manage all back-end data protection and high-availability functions.

Figure 32 illustrates the EqualLogic Unified Storage Architecture.

- It is a unified solution because the system is providing both block and file level storage.
- The storage pool shows a single pool that contains a NAS Reserve (managed by the NAS Appliance) plus other volumes that are providing block I/O storage targets not used by the FS NAS appliance. To emphasize this flexibility, we show a group of servers connected to the NAS Appliance through the LAN Switches (acting as CIFS/NFS clients) and connecting directly to the iSCSI SAN switches for block I/O.
- The EqualLogic Group Manager provides a single user interface for configuring and managing PS Series and FS Series functions.

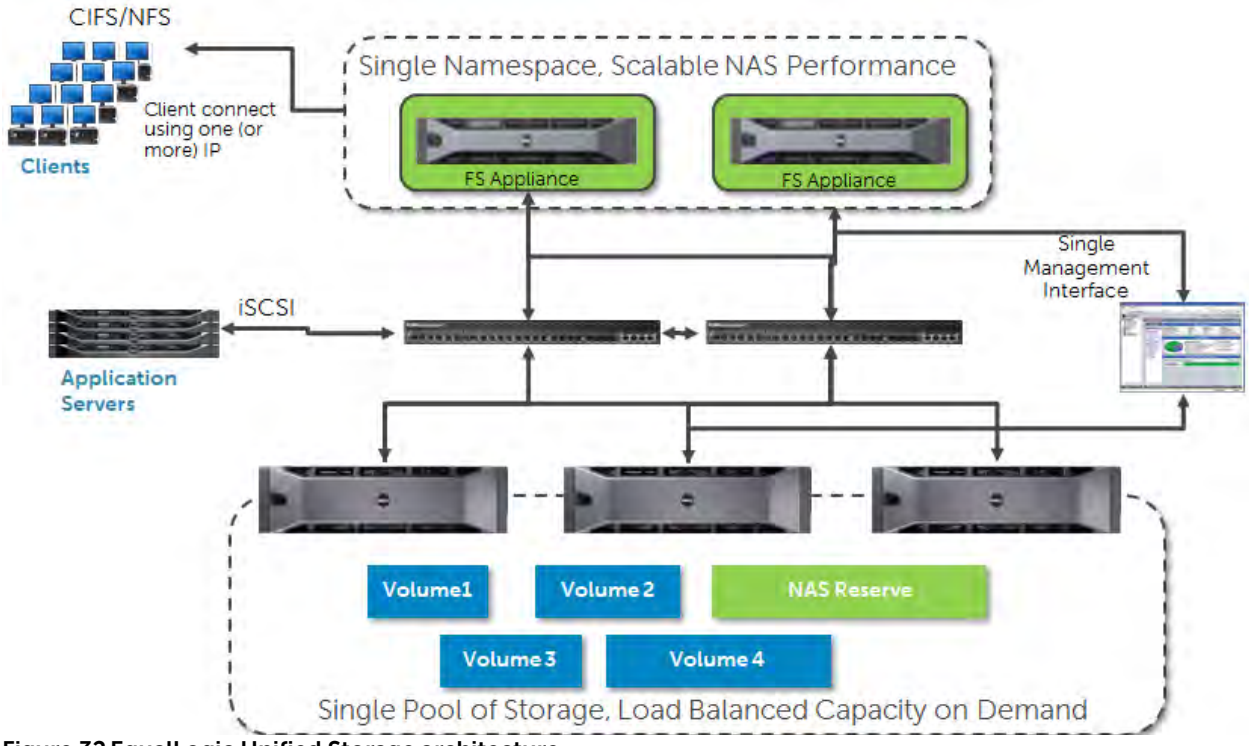


Figure 32 EqualLogic Unified Storage architecture

10.1.1 FS Series solution for file only storage

In a file only scenario the initiators on the FS Series appliance are the only iSCSI clients connecting to the PS Series arrays. The pools and volumes in the arrays provide storage for FS Series appliance file I/O only. This scenario is shown in Figure 33.

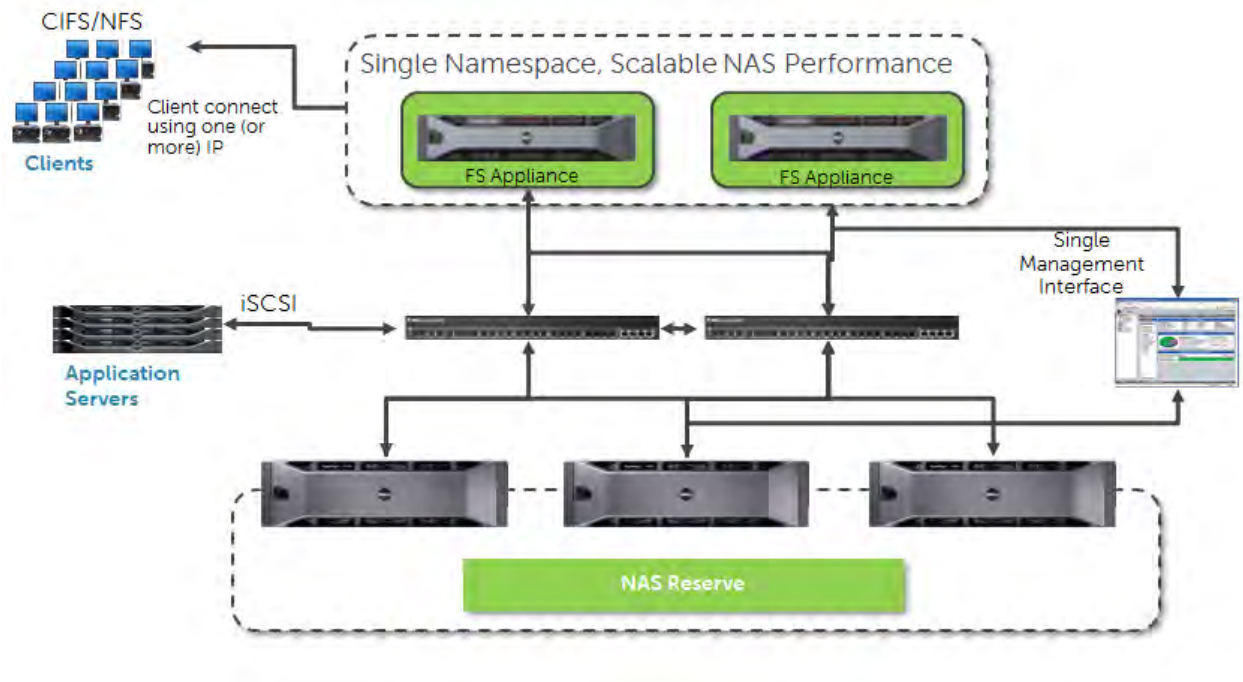


Figure 33 FS Series NAS (file only)

10.2 Dell FluidFS

Dell EqualLogic FS Series appliances, coupled with PS Series arrays, offer a high performance, high availability, scalable NAS solution.

FS Series Firmware V2.0 adds the following features:

- NAS Container Replication
- NAS Antivirus Service for CIFS Shares
- New Monitoring Features
- Support for FS7600 and FS7610 Appliances

Dell FluidFS[3] is a high performance clustered file system. It provides fully interoperable multi-protocol file sharing for UNIX, Linux, and Windows® clients using standard CIFS and NFS file access protocols and authentication methods (Active Directory, LDAP, NIS). Some of the key features provided in FluidFS include:

Single namespace: FluidFS presents a single file system namespace (up to 509TB usable capacity), accessible via a virtual IP address, regardless of cluster size.

High availability and reliability: The FS Series architecture includes a high speed cluster interconnect between controller nodes, write cache mirroring, failsafe journaling, and data integrity checks. In a FluidFS based FS Series cluster, a controller node can fail without affecting data availability or causing data loss, even if write operations were occurring at the time of failure.

[3] FluidFS technical whitepaper: http://www.dellstorage.com/WorkArea/DownloadAsset.aspx?id=1578

Snapshots: FluidFS snapshots are read only and redirect-on-write. They are created and managed by the FS Series appliance to provide file system level snapshot capability. They function independently of and have no impact on the operation of PS Series array based snapshots.

> **Note:** FS Series FluidFS snapshots and PS Series volume based snapshots function independently and have no impact on each other.
> Please see the following whitepaper for more information on Dell FluidFS:
> Dell Fluid File System:
> http://www.dellstorage.com/WorkArea/DownloadAsset.aspx?id=1578

NAS Replication: NAS replication is point-in-time and uses file system snapshot technology. NAS replication can be configured on a per NAS container basis.

11 FS Series NAS Appliances

The FS7500 is the premier offering in the Dell FS Series product line. Table 45 lists the basic functional details for each FS Series product.

Table 45 FS Series Models

FS Series Model	System Components	I/O Connections
FS7500	2 x 1U Intel/Linux based NAS appliance; dual active-active controller configuration 1 x 1U Battery backup power supply unit (BPS)	**Client LAN:** 8x1GbE per appliance (4 per controller node) **iSCSI SAN:** 8x1GbE per system (4 per controller node) **Controller Interconnect:** 8x1GbE per system (4 per controller node)
FS7600	2U NAS appliance with two active/active NAS controllers. 2 x Intel 5620 Quad Core processors per NAS controller 24GB memory per NAS controller Redundant hot pluggable power supples and fans	Client LAN: 8x1GbE per appliance (4 per controller node) iSCSI SAN: 8x1GbE per system (4 per controller node)
FS7610	2U NAS appliance with two active/active NAS controllers. 2 x Intel 5620 Quad Core processors per NAS controller 24GB memory per NAS controller Redundant hot pluggable power supples and fans	Client LAN: 4x10GbE per appliance (2 per controller node) iSCSI SAN: 4x1G0bE per system (2 per controller node)

11.1 Equallogic NAS appliance supported configuration limits

The features and supported configuration limits for the NAS system are provided in Table 46.

Table 46 NAS cluster configuration limits

Attribute	Single NAS appliance cluster (Two NAS controllers)	Dual NAS appliance cluster (Four NAS controllers)
Maximum NAS reserve	509 TB usable	509 TB usable
Minimum NAS reserve	512 GB	1024 GB
Maximum file size	4 TB	4 TB
Maximum local groups	300	300
Files in a cluster	64 billion	128 billion
Directories in a cluster	34 billion	68 billion
Containers in a cluster	256	512
Minimum container size	20 MB	20 MB
Maximum container size	Available NAS Reserve	Available NAS Reserve
Snapshots in a cluster	10,000	10,000
Snapshot schedules for a container	64	64
Snapshot schedules in a cluster	1024	1024
Snapshots per container	512	512
NFS exports in a cluster	1024	1024
CIFS shares in a cluster	1024	1024
CIFS active client connections in cluster	1500	3000
Local users in a cluster	300	300
Maximum quota rules per NAS container (user quotas)	512	512
Maximum quota rules per cluster (user quotas)	100,000	100,000
Maximum number of containers enabled for replication	100	100
Maximum number of active container replications in progress (rest are queued)	10	10

Attribute	Single NAS appliance cluster (Two NAS controllers)	Dual NAS appliance cluster (Four NAS controllers)
Maximum replication partners	16	16
Number of NAS controllers in replication source and destination clusters	Each NAS cluster in the replication partnership must contain the same number of NAS controllers.	Each NAS cluster in the replication partnership must contain the same number of NAS controllers.

The limits shown for files and directories, containers, NFS exports, CIFS shares, and container snapshots are the maximum supported limits. The group will not prevent you from exceeding a limit and will not warn you when you reach or exceed a limit.

11.2 Initial NAS cluster valid configurations

The table below describes the configurations allowed during initial configuration.

Table 47 Initial valid configurations

Cluster	Controller Pair 1	Controller Pair 2
Two-controller	FS7500	NA
Two-controller	FS7600	NA
Two-controller	FS7610	NA
Four-controller	FS7500	FS7500
Four-controller	FS7500	FS7600
Four-controller	FS7600	FS7600
Four-controller	FS7610	FS7610

Table 48 Valid add-controllers pairs

Controller Pair 1	Controller Pair 2
FS7500	FS7500
FS7500	FS7600
FS7600	FS7600
FS7600	FS7500
FS7610	FS7610

11.3 FS7500 system components

The system components in the initial release of the EqualLogic FS Series NAS appliance[4] (the FS7500) consist of two controller node units and one backup power supply (BPS) shelf containing two redundant power supplies. The system components and required power cabling paths are shown in Figure 34.

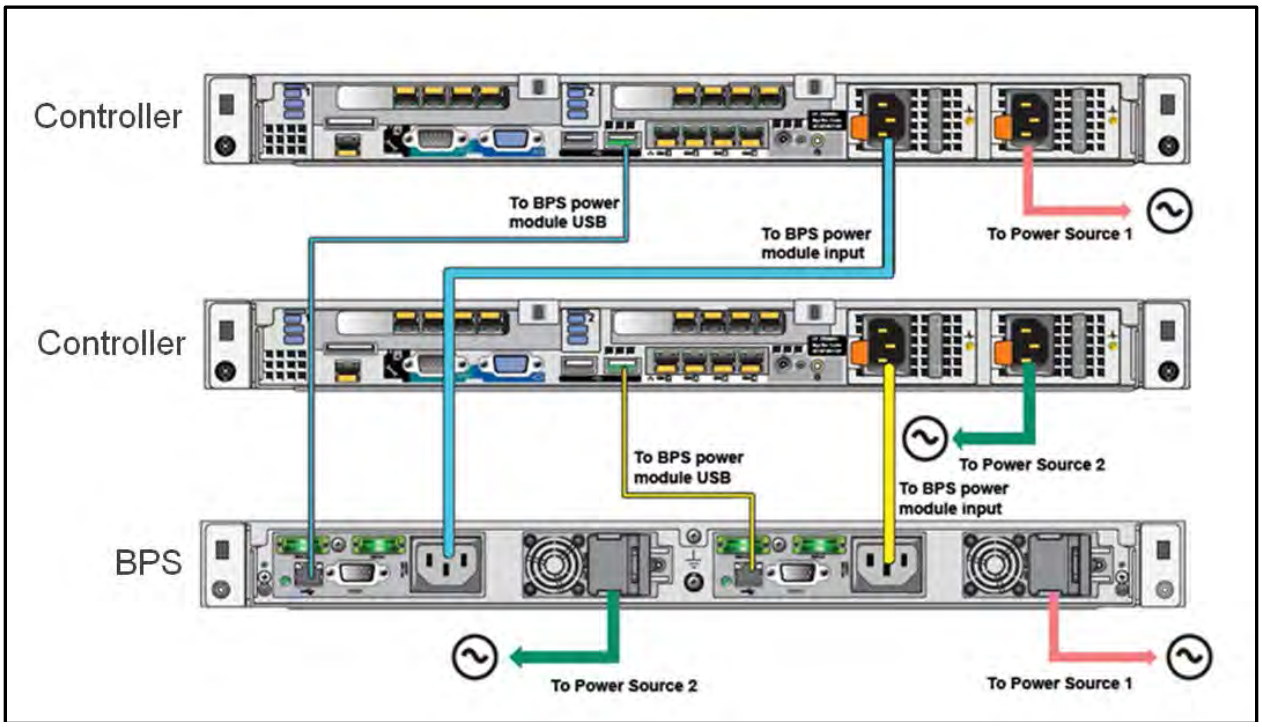


Figure 34 FS7500 system components and power cabling

[4] For detailed information on racking and power supply cabling see the *FS7500 Installation and Setup Guide*.

System controllers

The system contains dual active-active controller nodes with large onboard battery-backed caches and 24 GB of battery protected memory per node. They operate in an active-active environment mirroring the system's cache. Each node regularly monitors the BPS battery status. They require the BPS to maintain a minimum level of power stored during normal operation to ensure that they can execute a proper shutdown procedure in the event of power interruption.

Backup power supply

The BPS is designed to provide power necessary for controllers to execute a proper shutdown procedure. The BPS also enables the FS7500 controllers to use their onboard cache as NVRAM. This provides the clustered solution with enough time to write all the data from the cache to disk in the event a controller node experiences a loss of power.

11.4 FS7500 file system operation on controller failover

In the FS7500 system, writes to one controller node cache are mirrored to the other controller node cache before the write is acknowledged and committed. This prevents loss of any in-flight write I/O if a controller failure occurs. If a controller node failure occurs the cluster is put in journal mode. This triggers all I/O normally written to the mirror cache to be written to a journal file instead. Client load balancing in the FS7500 make this process transparent from a client point of view.

11.5 FS7600 components

The FS7600 and FS7610 NAS Appliance consist of two hot pluggable redundant NAS controllers per appliance, redundant hot pluggable power supplies, and fans. Each NAS controller contains a built-in battery backup to protect cache in case of controller power failure. The system components are shown in Figure 35 and Figure 36.

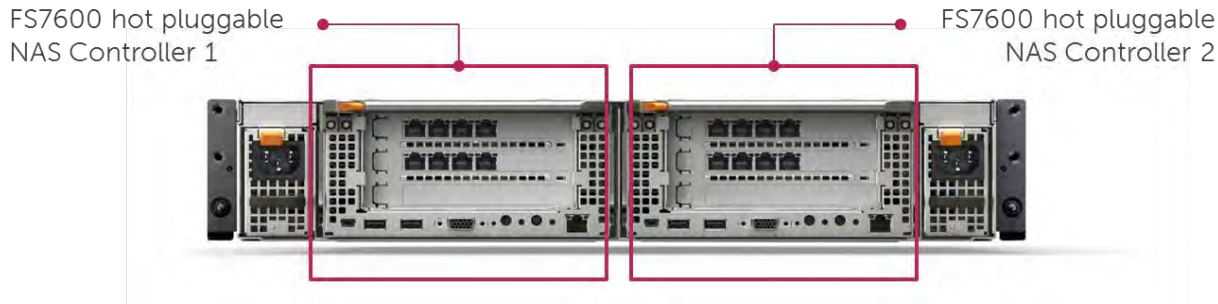


Figure 35 Rear view of the FS7600 NAS appliance

11.6 FS7610 components

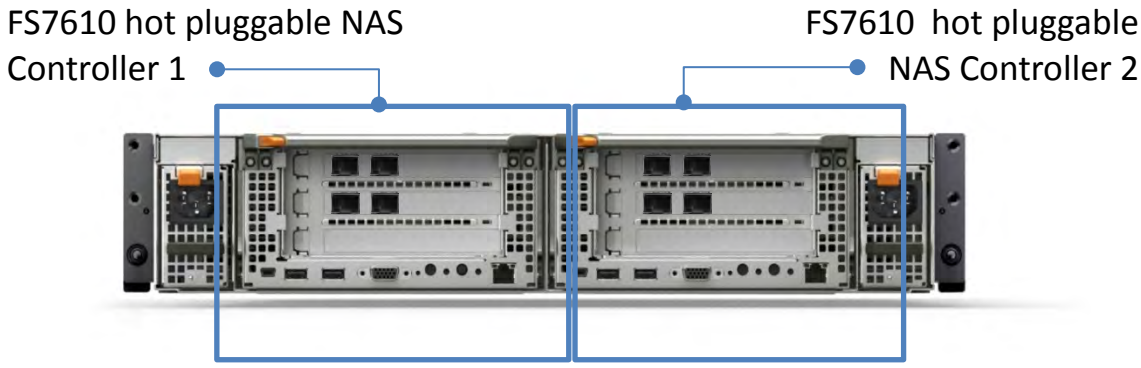


Figure 36 FS7610 NAS Appliance

12 FS Series file level operations

In this section we provide an overview of key FS Series Appliance features along with some operational limits. Please refer to the *NAS Service Management Tasks* section of the EqualLogic PS Series Group Administration Manual (version 5.1 or later) for detailed descriptions and all administrative task procedures.

12.1 NAS cluster

The NAS cluster is the logical run-time container in which one or more NAS containers are created. From the point of view of NAS clients, the NAS cluster is a virtual file server that hosts multiple CIFS shares or NFS exports. You can only have one NAS Cluster per EqualLogic group. NAS clients connect to file storage through a single NAS client virtual IP address. Client connections are load balanced across the available NAS controllers in the cluster. When you configure a NAS cluster, you specify the network configuration for the cluster and the amount of storage pool space consumed by the NAS Reserve.

12.2 NAS reserve

The NAS reserve is the storage space allocated to the NAS Cluster from an EqualLogic PS Series storage pool. The NAS reserve has the following properties:

- It resides within a single EqualLogic storage pool. Only one storage pool in the EqualLogic storage group can be used for allocating the NAS reserve.
- It is formatted with the Dell FluidFS.
- A fixed amount of storage space in the pool (512GB per controller node pair) is consumed by NAS cluster metadata. You should add 512GB per controller node pair to the calculated size of the NAS Reserve to compensate for this.
- The NAS reserve can be expanded later when more storage pool space is available.
- At the array level, the NAS reserve is comprised of a variable number of automatically created volumes (from a minimum of two volumes up to a maximum of 34). The actual number of storage volumes created within the pool depends on the reserve size setting.

Note: You cannot decrease the size of a NAS Reserve once it has been created.

12.2.1 Relationship between PS Series groups, pools and NAS reserve

The relationships are illustrated in Figure 37. As shown in the figure, the NAS Reserve pool can exist in a group hosting a single pool, or in a group hosting up to three other pools simultaneously.

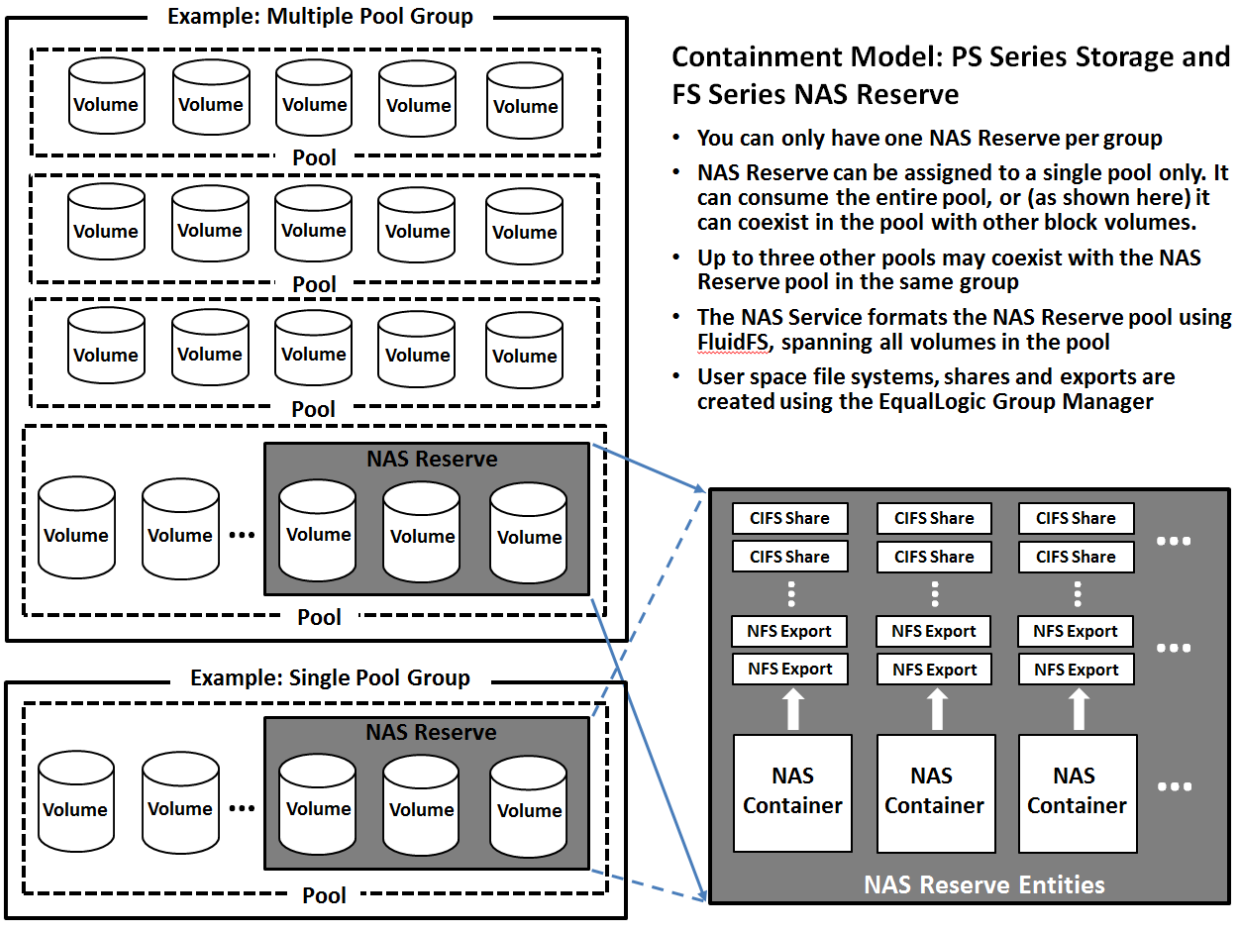

Example: Multiple Pool Group
Volume
Pool
NAS Reserve
Example: Single Pool Group
Containment Model: PS Series Storage and FS Series NAS Reserve
You can only have one NAS Reserve per group
NAS Reserve can be assigned to a single pool only. It can consume the entire pool, or (as shown here) it can coexist in the pool with other block volumes.
Up to three other pools may coexist with the NAS Reserve pool in the same group
The NAS Service formats the NAS Reserve pool using FluidFS, spanning all volumes in the pool
User space file systems, shares and exports are created using the EqualLogic Group Manager
CIFS Share
NFS Export
NAS Container
NAS Reserve Entities

Figure 37 Containment Model: PS Series Storage and FS Series NAS Reserve

The addition of the FS Series NAS appliance to a PS Series group does not change the functional behavior of PS Series groups and pools. PS Series groups and pools are explained in more detail in sections 5.1 and 5.1.1.

12.3 NAS Container

To provision NAS storage, you need to create NAS containers within the NAS cluster. Inside a NAS Container you can create multiple CIFS shares and/or NFS exports. Access to all shares and exports is through a single client virtual IP address.

You can create a single large container, or you can create many separate containers. You can apply different backup, snapshot, security, and quota policies to each container. Creating multiple containers gives you the flexibility to apply different management policies to different containers. The number and size of the containers in a NAS cluster depends on the storage needs of your NAS clients and applications. You can increase and decrease the size of containers as needed. The relationship between the NAS Cluster, NAS Reserve, NAS Containers, CIFS shares and NFS exports is shown in Figure 38 below. Note that you can setup "mixed protocol" access to portions of a container. This means that the same portion of the container can be simultaneously mounted by NFS clients and mapped by CIFs clients.

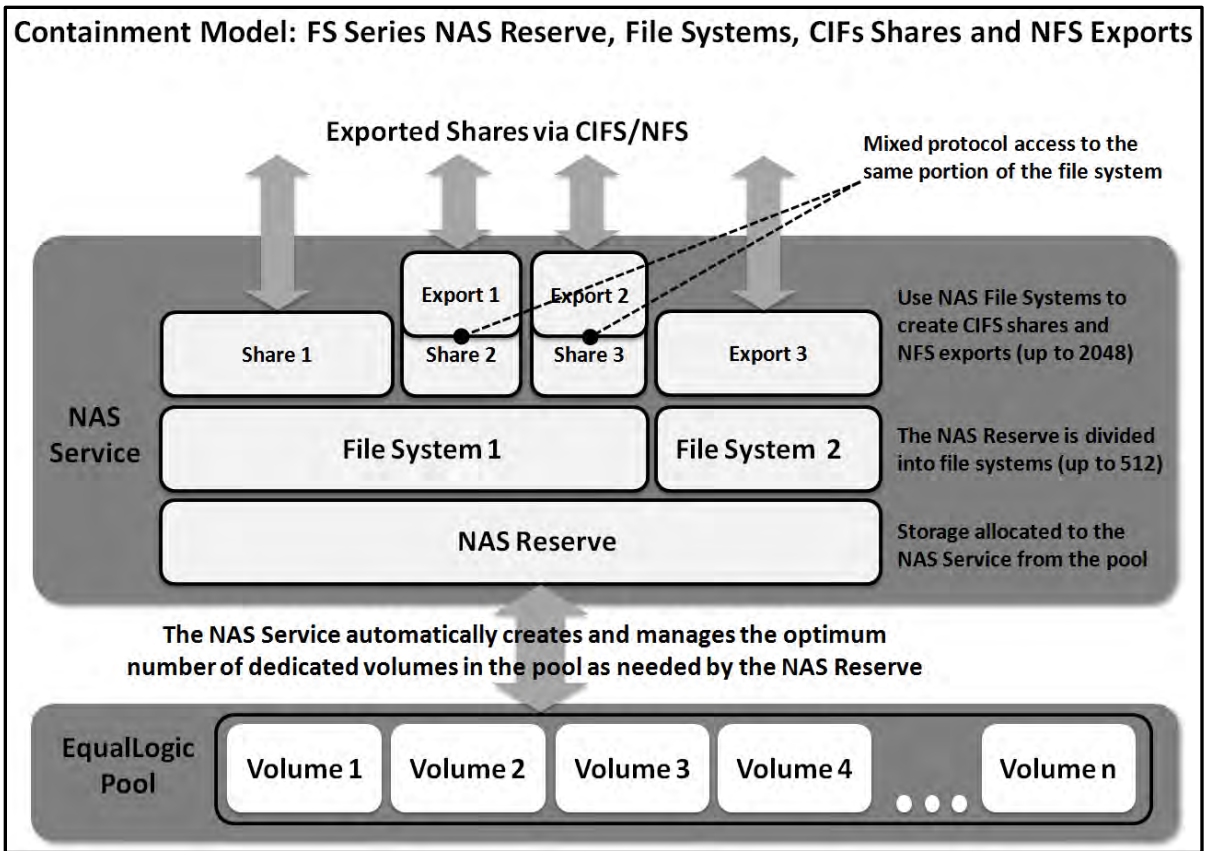

Containment Model: FS Series NAS Reserve, File Systems, CIFs Shares and NFS Exports
Exported Shares via CIFS/NFS
Mixed protocol access to the same portion of the file system
Export 1
Export 2
Share 1
Share 2
Share 3
Export 3
Use NAS File Systems to create CIFS shares and NFS exports (up to 2048)
NAS Service
File System 1
File System 2
The NAS Reserve is divided into file systems (up to 512)
NAS Reserve
Storage allocated to the NAS Service from the pool
The NAS Service automatically creates and manages the optimum number of dedicated volumes in the pool as needed by the NAS Reserve
EqualLogic Pool
Volume 1
Volume 2
Volume 3
Volume 4
Volume n

Figure 38 Containment Model: FS Series NAS Reserve, Containers, Shares and Exports

12.3.1 NAS Container security styles

There are three security style options that can be applied to a NAS Container:

- **UNIX**: Controls file access using UNIX permissions in all protocols. A client can change permissions only by using the chmod and chown commands on the NFS mount point. You can also specify the UNIX permissions for files and directories created in the container by Window clients. See About UNIX Permissions for Windows Directories and Files. With UNIX file security style, Windows clients cannot change any file or directory permissions. Read and write access is controlled by the UNIX permissions for Windows files and directories, which you set in Group Manager.
- **NTFS**: The default security style for NAS containers is NTFS. Controls file access by Windows permissions in all protocols. A client can change the permission and ownership by using the Windows Security tab. With NTFS file security style, all access permissions are controlled by the Windows administrator by using access control lists or share level permissions.
- **Mixed**: Supports both NTFS and UNIX security styles. The permissions and ownership for a file or directory will be the last one set. Permissions and access rights from one protocol to another are automatically translated. Mixed is a combination of NTFS and UNIX security style. A Windows user can override UNIX user settings, and a UNIX user can override Windows user settings. This security style is only recommend for scratch storage where maintaining permissions is not important.

The default security style for NAS containers is NTFS. Referring to **File System 1** in Figure 38, you could assign any of the three container security styles to it. Given that portions of it are simultaneously

accessible as CIFs shares and NFS exports, a general rule of thumb for how to assign the security style would be as follows:

- If your users are predominantly Linux/UNIX based, use the UNIX style.
- Likewise, if your users are predominantly Windows/CIFS based, use the NTFS style.
- If you have mixed clients, use the style applicable to majority of users. Then create a user mapping of your Windows to Linux users or vice versa. The user permissions are equivalent to the mapped user. Note that the Mixed Mode container security style should not be used for cases where you have a mix of both CIFS and NFS clients.

12.4 NAS Container snapshots

Dell FluidFS snapshots are created and managed by the FS Series appliance to provide a container level snapshot capability. They are read-only and are created using a redirect-on-write method. This method is also referred to as allocate-on-write. This approach requires only one I/O operation and delivers higher write performance.

> **Note:** FS Series FluidFS snapshots and PS Series volume based snapshots function independently and have no impact on each other.
> Please see the following whitepaper for more information on Dell FluidFS snapshot behavior:
> **Dell Fluid File System:**
> http://www.dellstorage.com/WorkArea/DownloadAsset.aspx?id=1578

Additional FluidFS snapshot limits and considerations:

- Each NAS Container has its own snapshot policy.
- The minimum unit a snapshot can act on is a NAS Container.
- The maximum number of snapshots you can retain per container is 512. The maximum snapshots retained per FS appliance cluster is 10,000.
- When you create a snapshot, all shares and exports within that container are included. If a particular share or export will require frequent snapshots, you should group it with others that have the same requirement, or dedicate a container to that particular share or export.
- The FluidFS snapshot implementation allows for end users to individually select and restore previous versions of files. The .snapshot directory contains all the snapshots taken for the container. Browse to the correct snapshot directory to access the files needed to be restored to the share. Windows users can also use the "Previous Version" feature of Windows files and directories to restore individual files or entire folders.

Sizing limits and space utilization considerations for FluidFS snapshots:

- The size of the snapshot reserve is specified as a percent of the container size. It is set at container creation and changes according to the % reserve setting whenever the size of the container is changed.
- The default snapshot reserve size is 50%. It is common to select a reserve size in the range of 25%. The maximum snapshot reserve size is 90% of the FS size.
- The snapshot reserve capacity is counted against the NAS container reserve space. However, the snapshot reserve space is not enforced. Container data can fill the snapshot reserve space.

- Snapshot reserve space utilization is a function of the data change rate in the container.
- Old snapshots are deleted to make room for new snapshots if enough snapshot reserve is not available.
- NDMP based backup automatically creates a snapshot, from which the backup is created.

12.5 NAS Snapshots and Replication

The FS76X0 further ensures business continuity through support for file system, point-in-time snapshots and snapshot-based, asynchronous replication. It is important to note that Group Manager block-level replication cannot be used as a data protection strategy for FS76X0 data.

12.5.1 Snapshots

Fluid File System leverages redirect-on-write snapshots that only require one I/O per write operation and thus avoid the performance degradation of the more traditional copy-on- write approach. Fluid File System freezes an image and maintains a point-in-time backup of the data. Each individual NAS container can have its own snapshot policies. The snapshots can be recovered instantly by administrators and are also available to end-users through the NAS protocols as read-only views.

12.5.2 Replication

The FS76X0 and FS7500 (using FluidFS series firmware 2.0) now support snapshot-based, asynchronous NAS replication to or from another FS7500 or FS76X0 system in a different EqualLogic Group at a peer (local or remote) location. An EqualLogic Group containing FS appliances forms a partnership with another EqualLogic group to replicate data across the two groups.

The NAS replication uses snapshots to replicate only the changed blocks across sites. NAS replication works at the NAS container level and can be scheduled (hourly, daily, weekly, etc.) per NAS container. The replication site must have the free capacity in the NAS reserve to replicate the NAS containers chosen for replication. All NAS container snapshots on the primary are also replicated to the secondary site.

> **Note:** NAS replication uses the SAN side network to replicate file data. Refer to the NAS Network Setup Guide for details on making you're the customer network is ready to support NAS replication.

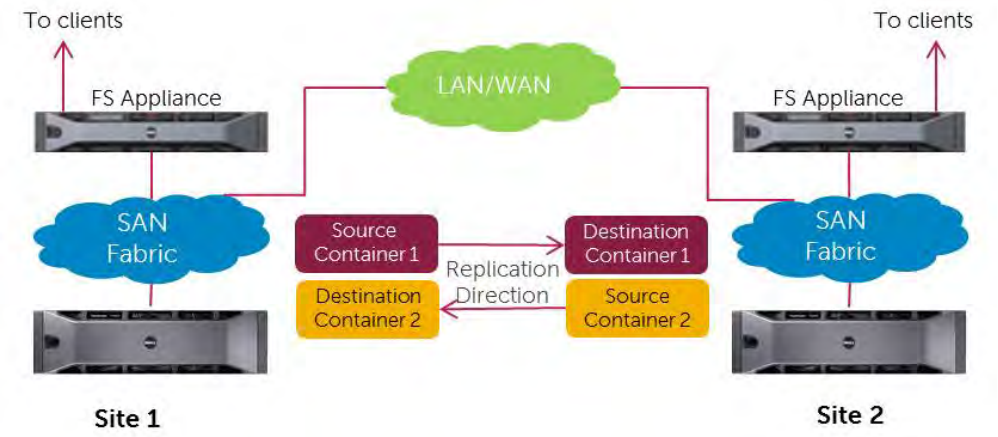


Figure 39 NAS replication

Replication requirements:

- Both sites must be EqualLogic FS appliances. Replication from FS7600 to any other Dell product using FluidFS is not supported.
- Each site (or EqualLogic Group) must have the same number of NAS controllers or NAS appliances.
- FS appliances must run firmware version 2.0 or higher and PS arrays must run version 6.0 or higher.
- There must be a network link between the two sites (or EqualLogic groups) via the SAN side network.
- If the client and SAN networks do not have a route to each other, a static route needs to be configured on FS NAS appliances on both the sites. Refer to NAS Replication Network Setup Guide for details on configuring network for NAS replication.

Important notes:

- You can replicate between FS7500 and FS76x0 as long as all FS appliances are running version 2.0 of FS Series Firmware (FluidFS).
- Replication of NAS containers within the same EqualLogic group or NAS cluster is not supported by FS appliances.
- One EqualLogic group can form replication partnerships with 16 other groups; however one NAS container in a group can be replicated to only a single replication partner.
- NAS replication operates at block level. After the first complete data replication, only changed blocks of a file and not the whole file (if any changes are made to a file) is replicated.
- There is no Manual Transfer Utility (MTU) available for NAS replication.
- The NAS replica is consistent to the last successful replication. If a you want multiple point in time copies of the NAS container at the replica site, snapshots should be created on the primary side. All the snapshots and the NAS container point in time copies will be available at the secondary site.
- The NAS replication does not use any replication reserve space on the pool.

Accessing data at the replication site:

- Read Only Access: A NAS replica can be made available to clients on secondary site by enabling Read Only access

- Temporary Promotion: Promote Read/Write with an ability to demote. However all writes are lost when container is demoted to resume replication from primary site.
- Permanent Promote: Promote Read/Write in case primary site or container is not available and clients need to be failed over to secondary site. All writes are preserved

13 FS Series NAS Configuration

In Section 11.3 we presented a high level view of the FS7500 NAS component architecture (See Figure 34). In this section we provide detailed connection diagrams demonstrating how to setup fully connected iSCSI SAN and client LAN connection paths for the FS7500 and FS7600/FS7610 appliances.

> **Note:** It is recommended to keep the client and SAN side networks physically separate and deploy two switches on both sides to provide redundancy in the event of a switch failure.

13.1 FS7500 connection paths

The FS7500 appliance is comprised of two peer system controller nodes. In the required fully cabled configuration, each controller node requires thirteen separate Ethernet cable connections. Thus a single FS7500 appliance comprised of two controller nodes requires a total of 26 Ethernet connections (four connecting to the client LAN switches and nine connecting to the iSCSI SAN fabric switches.)

Figure 40 shows the client LAN connection paths.

> **Note:** While it is possible to operate an FS7500 appliance in a partially cabled configuration, this configuration is not supported by Dell. You should use a fully cabled configuration in a production environment. You will also need to provision the required switch port count on the iSCSI SAN and client LAN sides of the system to support a fully connected configuration.

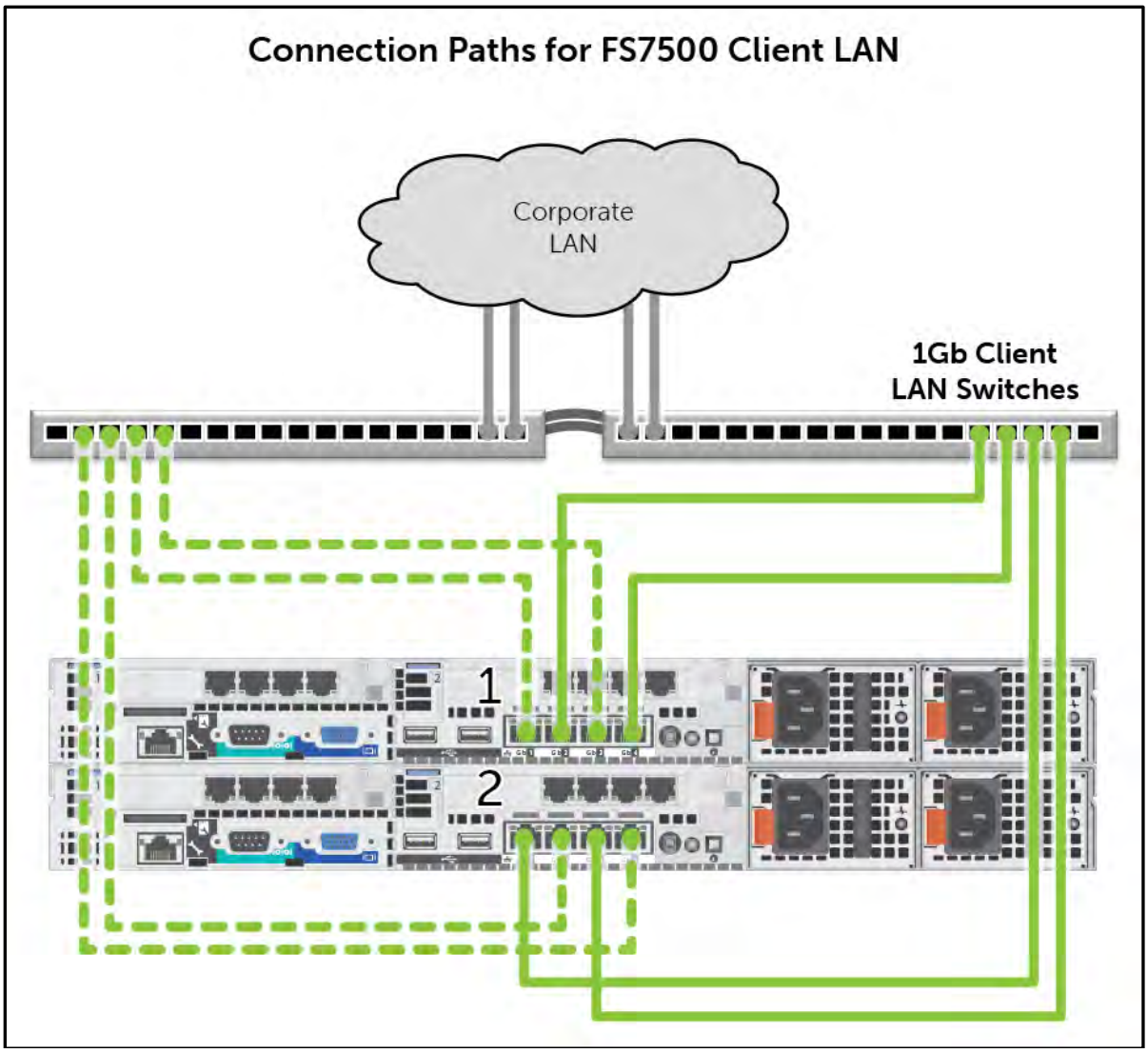


Figure 40 Connection Paths for FS7500 Client LAN

Figure 41 below shows the iSCSI SAN, IPMI, and node interconnect paths. Pay careful attention to how the controller ports alternate between redundant switch paths.

> **Note:** With the exception of the IPMI connection paths, corresponding ports on each controller node must connect to the same SAN switch. This connection pattern is shown in Figure 41.

Sizing the iSCSI SAN Inter-Switch Connection for FS7500

The inter-switch connection sizing guidelines provided in Section 7.2.2 also apply to FS7500 SAN design. The FS7500 mirrors the write cache between controller nodes. To accomplish this, all write operations are transmitted across the controller interconnect. Thus, it is very important that you follow connection pattern shown in Figure 41 to ensure corresponding ports are connected to the same switch. This connection pattern prevents controller interconnect traffic from crossing the inter-switch connection.

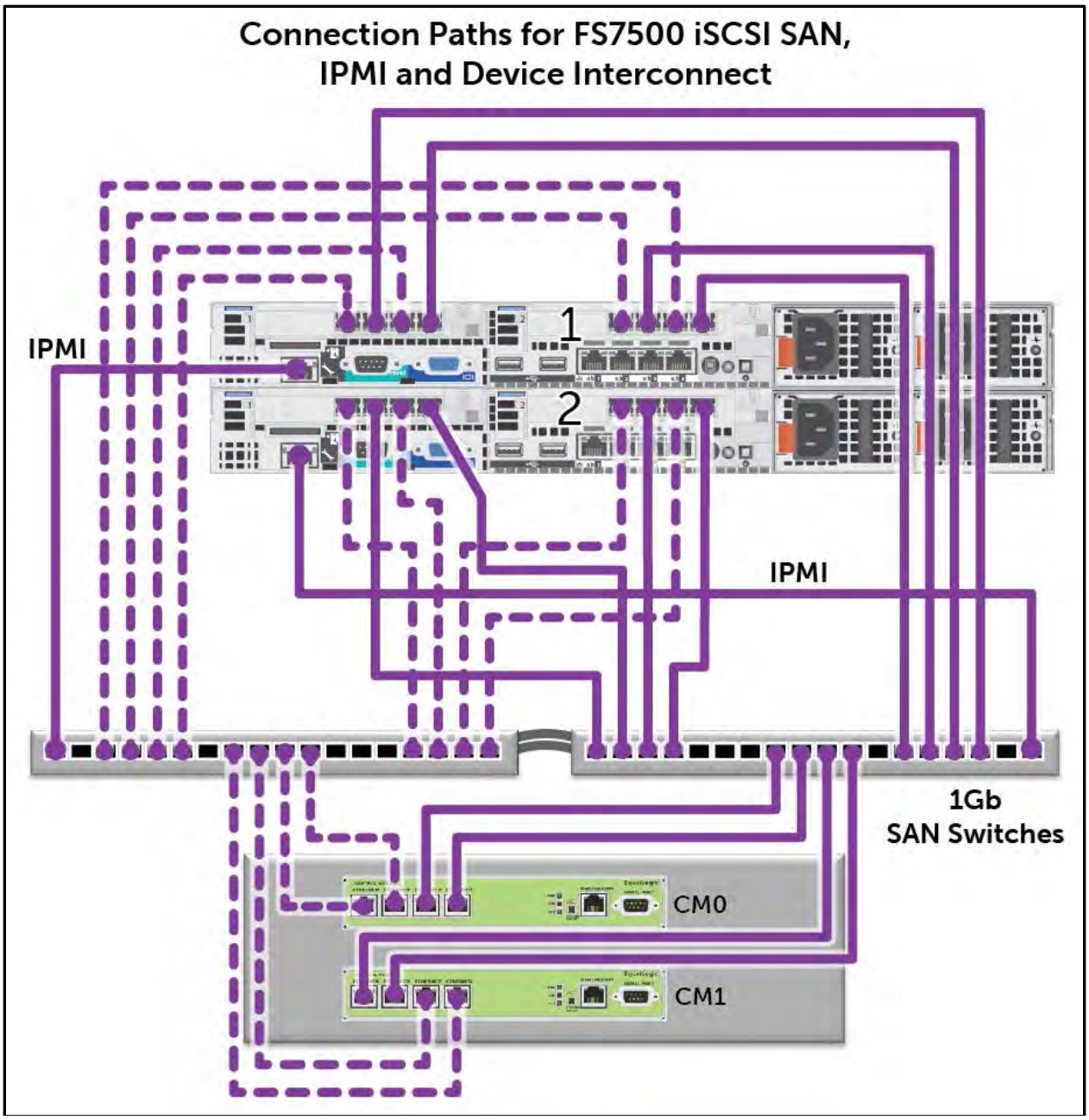


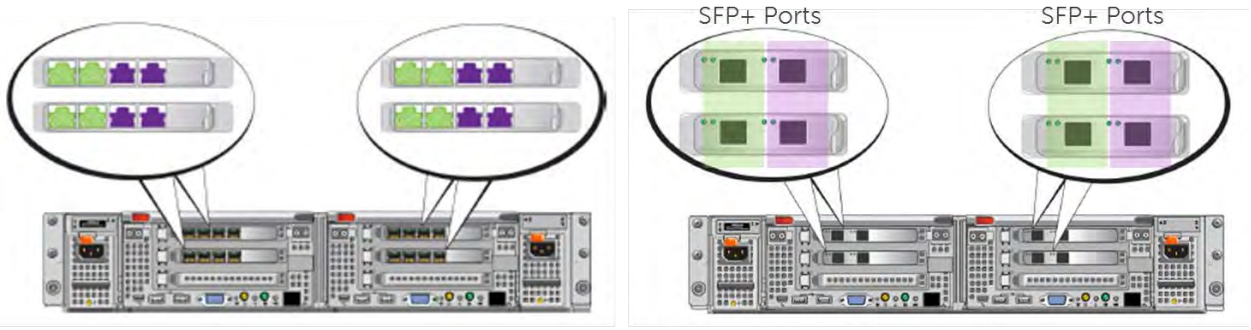
Figure 41 Connection Paths for FS7500 iSCSI SAN, IPMI and Controller Interconnect

13.2 FS7600/7610 connection paths

The Dell EqualLogic NAS appliances require the following networks:

- Client network: Used for client access to the NFS exports and CIFS shares hosted by the NAS cluster.
- SAN/internal network: Used for internal communication between the controllers and communication between the controllers and the EqualLogic PS Series SAN. The SAN and Internal networks use the same set of switches.

It is recommended to keep the client and SAN side networks physically separate and deploy two switches on both sides to protect against a switch failure.



Network	Network Usage	Ethernet Port Numbers
Client network	The Client network allows the clients (workstations, PCs, etc.) to access network shares.	Left port in all NICs
SAN / Internal networks	-The SAN network connections allow communication between the NAS cluster and the PS Series group (SAN). The SAN and internal networks use two different IP address ranges, but they exist on the same switch or VLAN. -The internal network connections allow communication between the two controllers.	Right port in all NICs

Figure 42 FS7600 AND FS7610 networks

See figures below for network connections.

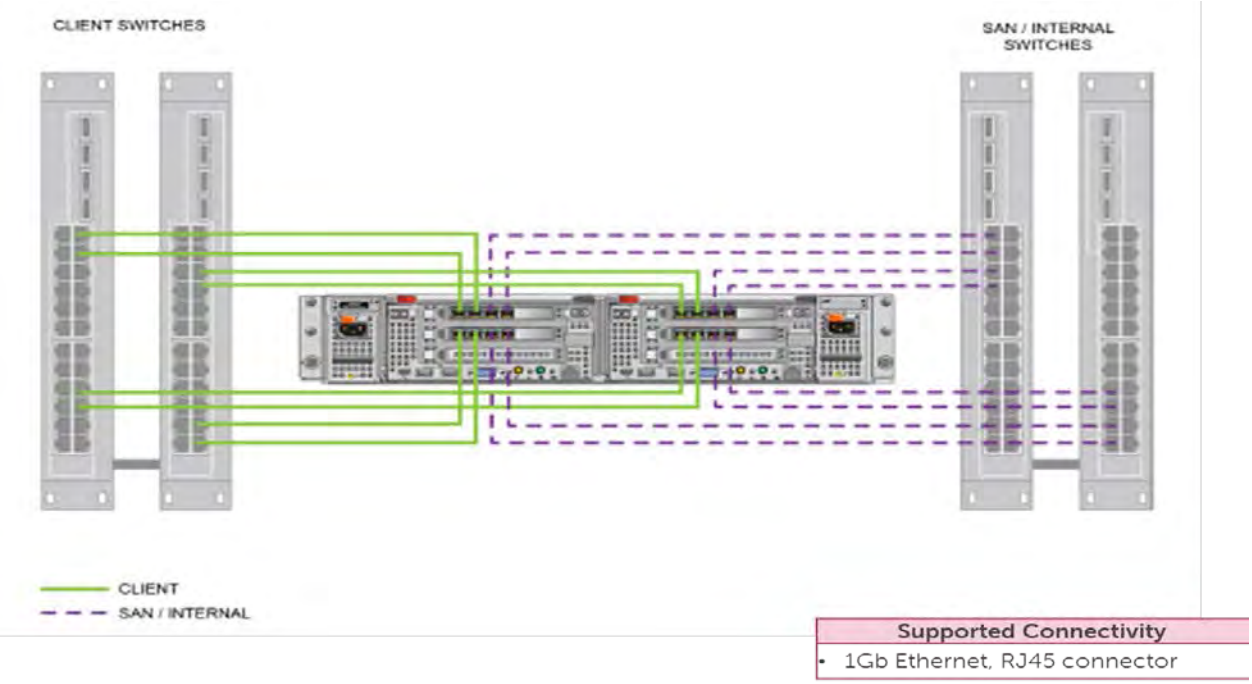


Figure 43 FS7600 network

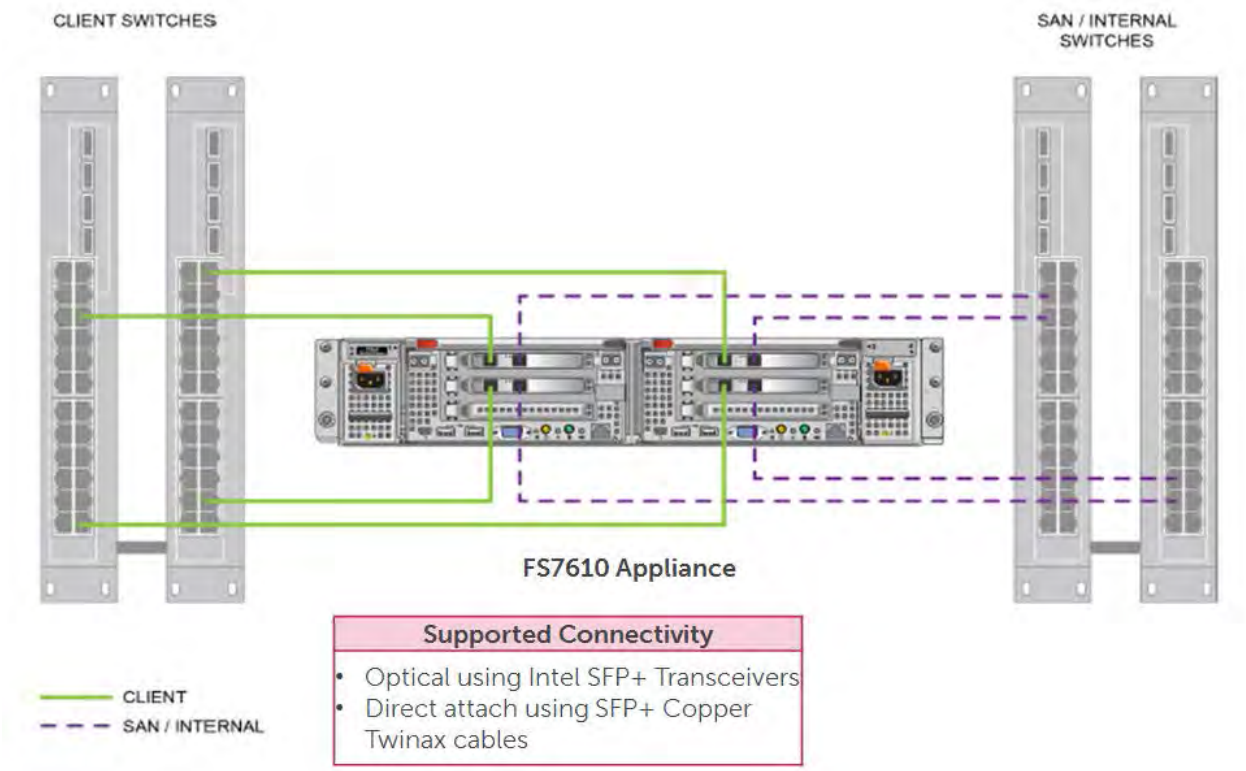


Figure 44 FS7610 network

Installation/Expansion

- If installing FS7500/FS76x0 into an existing EqualLogic SAN, verify the existing customer network meets the minimum requirements. Refer to FS76x0 installation guide for more information on network requirements. Early validation helps avoid issues during and after the install.
- FS7500/FS76x0 installation service is mandatory.
- All NAS appliances in a NAS cluster need to have either 1Gb or 10Gb connectivity. Appliances with different connectivity cannot be mixed in a NAS cluster.
    - FS7500 cluster can be expanded by using an FS7500 or FS7600 and NOT FS7610.
- There is no way to upgrade FS7500 to 10Gb Ethernet by replacing the network controllers or the NAS appliance. 10Gb functionality is provided by FS7610.

Network

- The two iDRAC IPMI ports on FS7500 are 100Mb. The SAN switches must be able to support 100Mb in addition to 1Gbps speed. Most 10Gb SFP+ switches (including 8024F and F10 S48) do not support 100Mb.
    - It is best to connect FS7500 to 1 Gb switches and then uplink 1Gb switches to 10Gb switches OR connect 100Mb iDRAC ports to 1Gb switches (that are connected to 10Gb switches) and FS7500 SAN/Internal to 10Gb switch ports that can run at 1Gb speed.
- FS appliances do not support client VIPs in multiple subnets and use of netgroups.

Data Protection

- FS76x0/FS7500 does not support replication to local NAS cluster
- EqualLogic block replication/clones/snapshots cannot be used to protect data managed by FS appliances. Use NAS replication to protect file data.

14 Data Center Bridging (DCB)

The enhancement to the Ethernet Specifications (IEEE 802.3 specifications) called Data Center Bridging (DCB) enables bandwidth allocation and lossless behavior for storage traffic when the same physical network infrastructure is shared between storage and other traffic.

The network is the fundamental resource that connects the assorted devices together to form the datacenter Ethernet infrastructure. These devices include the server hardware (along with the operating system and the applications that run on the host) and the storage systems that host application data. Sharing this Ethernet infrastructure with multiple traffic types (LAN and SAN) requires a fairness mechanism to provide bandwidth allocation and flow control for each type of traffic. Without such a mechanism, the Local Area Network (LAN) traffic and the Storage Area Network (SAN) traffic would have to be separated onto their own dedicated networks.

When the SAN and LAN networks are shared, all traffic is equal, unless Quality of Service (QoS) and/or Class of Service (CoS) is used.

Traditional or Non-DCB QoS (IEEE 802.1p) is not supported for EqualLogic implementations.

In a shared network environment, LAN and SAN traffic can impact each other and QoS may not solve this because of QoS implementation differences with vendors. These different implementations include:

- Number of Queues managed
- Relative priorities between queues
- Bandwidth reserved per queue

Also traditional QoS lacks selective flow control for each traffic type. Regular IEEE 802.3 PAUSE will pause the entire link and not selectively pause LAN or SAN traffic. This flow control ability is important for reducing congestion on switches and to enable fair sharing of resources between traffic types.

Other methods of network sharing:

- VLANs offer port-level security and segregation, but do not provide guaranteed bandwidth or quality of service.
- NIC partitioning (NPAR) manages traffic within the host. Once the network data exits the NIC to the switch, any QoS/bandwidth management enforced by NPAR is not honored by the switch.

> **Note:** It is not recommended to share network infrastructure without DCB. DCB is the only recommended method of converging SAN and LAN in an EqualLogic Storage environment.

14.1 DCB Overview

DCB is a collection of Standards designed to improve QoS networking and management in the Data Center that enables iSCSI SANs or FCoE SANs or both to converge with regular server LAN traffic on

the same physical infrastructure to increase operational efficiency, constrain costs, and ease network management.

There are primarily three progressive versions of DCB:

- Cisco, Intel, Nuova (CIN) DCBX
- Converged Enhanced Ethernet (CEE) DCBX or baseline DCBX
- Institute of Electrical and Electronic Engineers (IEEE) DCB.

DCB technologies based on standards include:

- PFC – Priority based Flow Control (802.1Qbb)
- ETS – Enhanced Transmission Selection (802.1Qaz)
- CN – Congestion Notification (802.1Qau)
- DCBx – Data Center Bridging Capability eXchange
  - Support for iSCSI application protocol priority with DCBX (also known as iSCSI TLV)

> **Note:** DCB uses 10Gb and higher Ethernet only.

I/O convergence using DCB for Dell EqualLogic iSCSI storage is the future direction for "Converged iSCSI" in a lossless Ethernet environment.

Without DCB, it is typical to dedicate a separate physical network infrastructure for SAN traffic to guarantee the infrastructure for bandwidth and performance as shown in the figure below. For iSCSI, it is a best practice to dedicate a SAN comprised of Ethernet switches for server and storage inter-connection. In this case, the server and storage systems are configured with dedicated NICs to communicate on the iSCSI SAN.

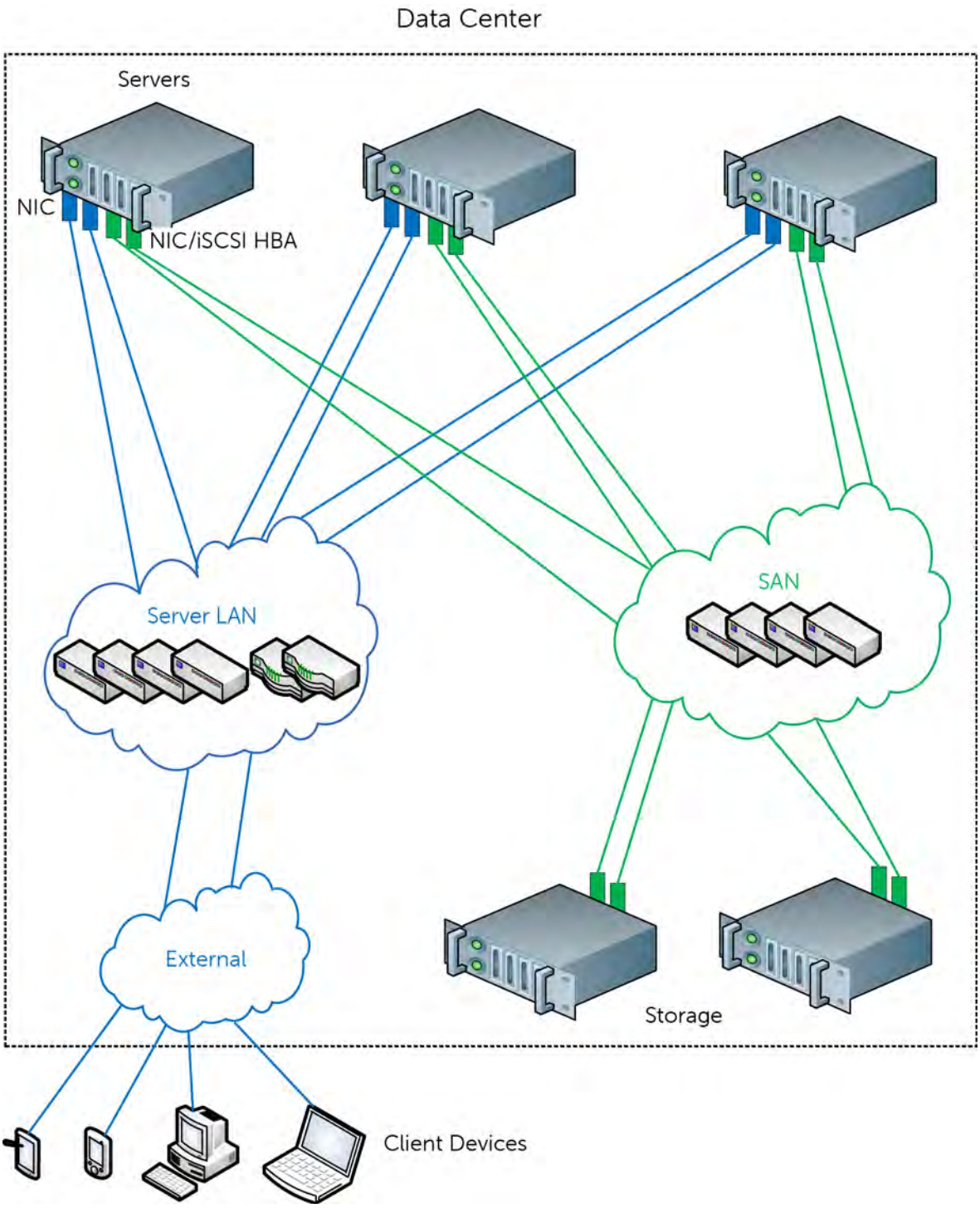


Figure 45 Physically separate dedicated network infrastructures

A converged network includes carrying both SAN and other network traffic such as server LAN on a single network infrastructure as shown in Figure 46. iSCSI can be converged with non-storage based server LAN traffic, allowing the network ports and the inter-connecting links to carry multiple traffic types or protocols. A DCB enabled converged Ethernet infrastructure includes the NIC/CNA on the end-devices (servers and storage arrays) along with the switching infrastructure.

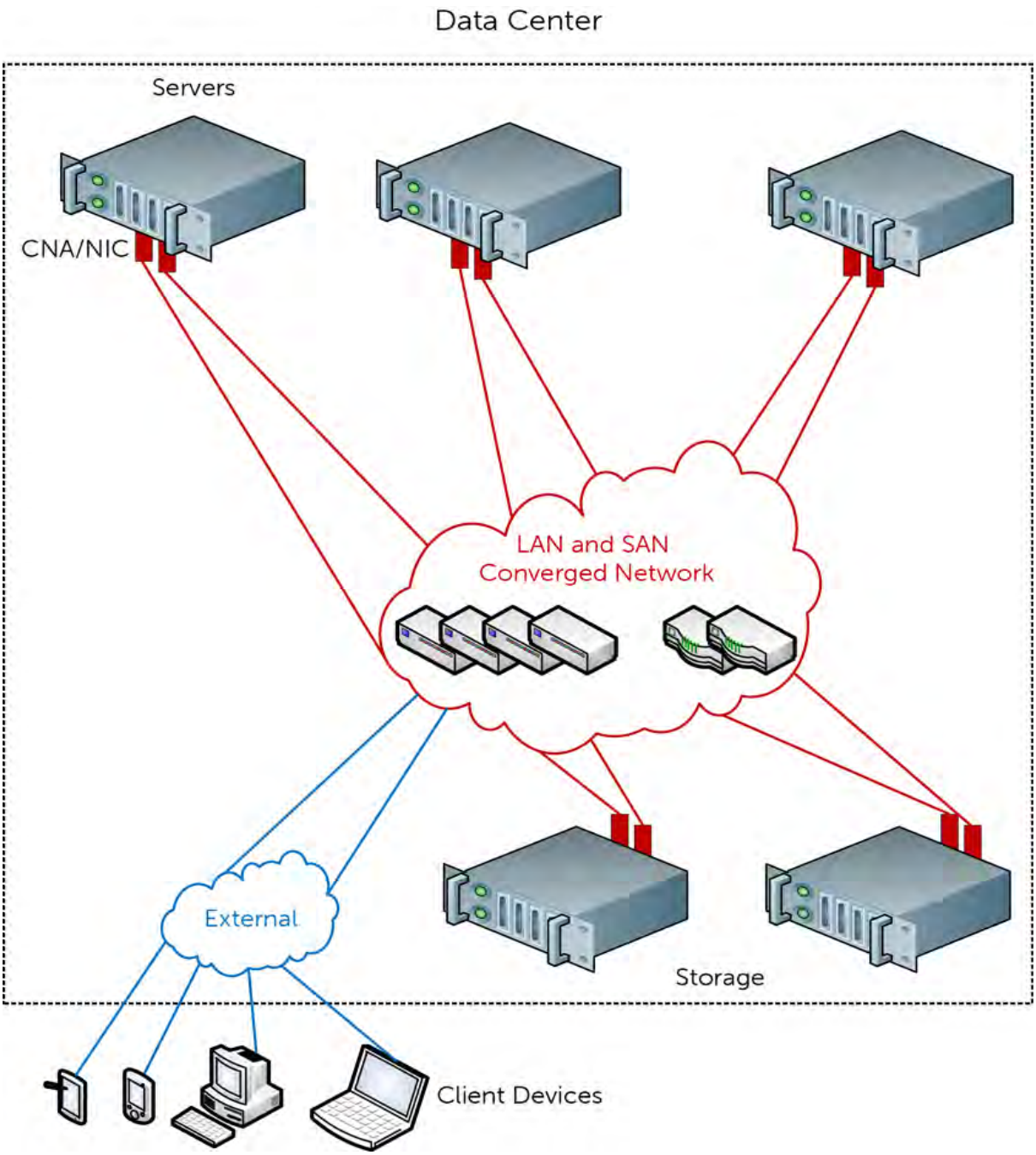


Figure 46 DCB enabled converged network infrastructure

14.2 DCB requirements for EqualLogic

It is required that all devices in the EqualLogic SAN support DCB for iSCSI when this functionality is enabled. If any device in the SAN does not support DCB, then DCB needs to be disabled at the switches for the entire SAN. Once all devices in the SAN are DCB compliant, then DCB can be re-enabled. Switches and server CNAs/NICs that are designated as DCB Supported in the EqualLogic Compatibility Matrix have been fully validated by Dell to ensure compatibility for EqualLogic SANs.

The minimum switch and server CNA/NIC requirements to support an end-to-end DCB solution with EqualLogic are:

- Data Center Bridging Exchange (DCBx) -DCB protocol that performs discovery, configuration, and mismatch resolution using Link Layer Discovery Protocol (LLDP)
- Application Priority (iSCSI TLV) - Switches must support configuration of a priority value for iSCSI protocol and advertisement to peer ports. Server NICs/CNAs: Must support iSCSI protocol in application priority (learned from the switch) and must support tagging outgoing iSCSI frames with this priority.
- ETS - Requires a dedicated traffic class or priority group for iSCSI priority
- PFC - Requires enabling PFC (no drop or lossless behavior) for iSCSI priority
  - Switches: Configure dedicated TC/PG for iSCSI priority with allocated bandwidth and PFC enabled for iSCSI priority
  - Server NICs/CNAs: Adhere to TC/PG mapping for iSCSI priority and PFC for iSCSI priority (learned from the switch)

Designing a converged network deployment with components that have no DCB support or partial DCB support is not recommended for end-to-end converged I/O. Configurations with components that have no DCB support or partial DCB support may result in loss of expected functionality or may be functional without meeting the converged network objectives.

> **Note:** It is not recommended to share network infrastructure with DCB and Non-DCB components.

For a complete list of components that support the DCB standards, see the: EqualLogic Compatibility Matrix

Additionally, there is a partner verified support program (PVSP) between Dell and VMWare. This agreement is to support Internet SCSI (iSCSI) over Data Center Bridging (DCB) Solution for ESX in a DCB environment with EqualLogic storage.

VMWare PVSP on VMWare website: EqualLogic SAN over DCB Solution for VMware ESXi 5.1

> **Note:** It is important to verify that all components in the SAN are listed in the EqualLogic Compatibility Matrix as DCB Supported, or that the components support all EqualLogic requirements for DCB.

14.3 Methods for configuring DCB

Configuration propagation: If this mechanism is supported by the switch model, then peer switches, downstream switches, and I/O Aggregator modules should accept the DCBX configuration from the source switches through certain Inter Switch Link (ISL) ports set as upstream ports capable of receiving DCBX configuration from peers (enabled with willing mode turned on the ports).

Manual configuration: If the internal propagation method is not supported by your switch model, then peer and intermediate layer switches must be manually configured with the same DCB parameters as the source switches with willing mode off across all ports.

The recommended operational mode is that switches use non-willing DCBx mode, while server NIC/CNAs and Storage ports operate in willing mode. This is the default behavior on most switches and the NICs/CNAs.

The DCB parameters (ETS, PFC, iSCSI application priority) are then configured in the switch devices and learned by the end-devices.

14.4 Basic Deployment Topology Example

This topology is an example topology of a single layer switching with rack servers.

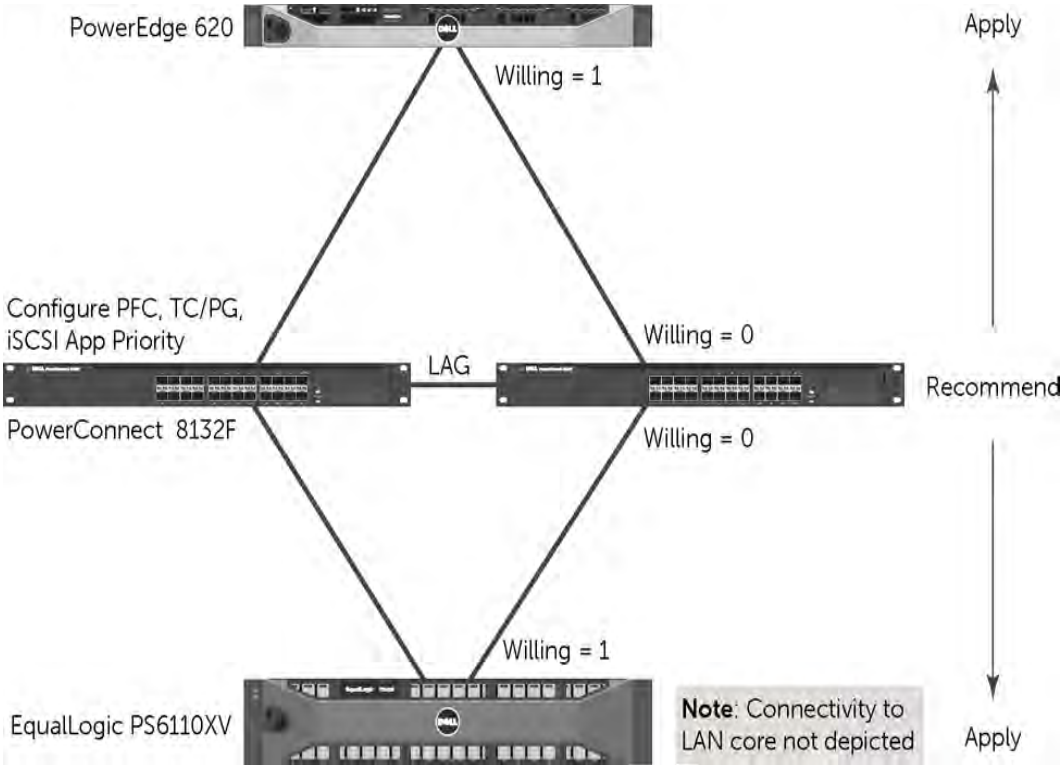


Figure 47 Single Layer Switching example with rack servers

PowerConnect 8132F supports the auto-up port role for learning DCBX configuration and the auto-down port role for internally propagating to a downstream device. Similary, Force10 S4810 supports the auto-upstream and auto-downstream port roles.

14.5 Tested SAN designs

This section shows three different end-to-end DCB SAN designs that were tested using the M1000e blade chassis. The topologies are tested deployments that can be used with EqualLogic storage arrays.

There are three categories of SAN designs for M1000e blade chassis integration:

- Blade IOM switch only
- ToR switch only

- Blade IOM switch with ToR switch.

14.5.1 Blade IOM switch only

Network ports of both the hosts and storage are connected to the M1000e blade IOM switches. No ToR switches are required. The switch interconnect can be a stack or a LAG, and no uplink is required.

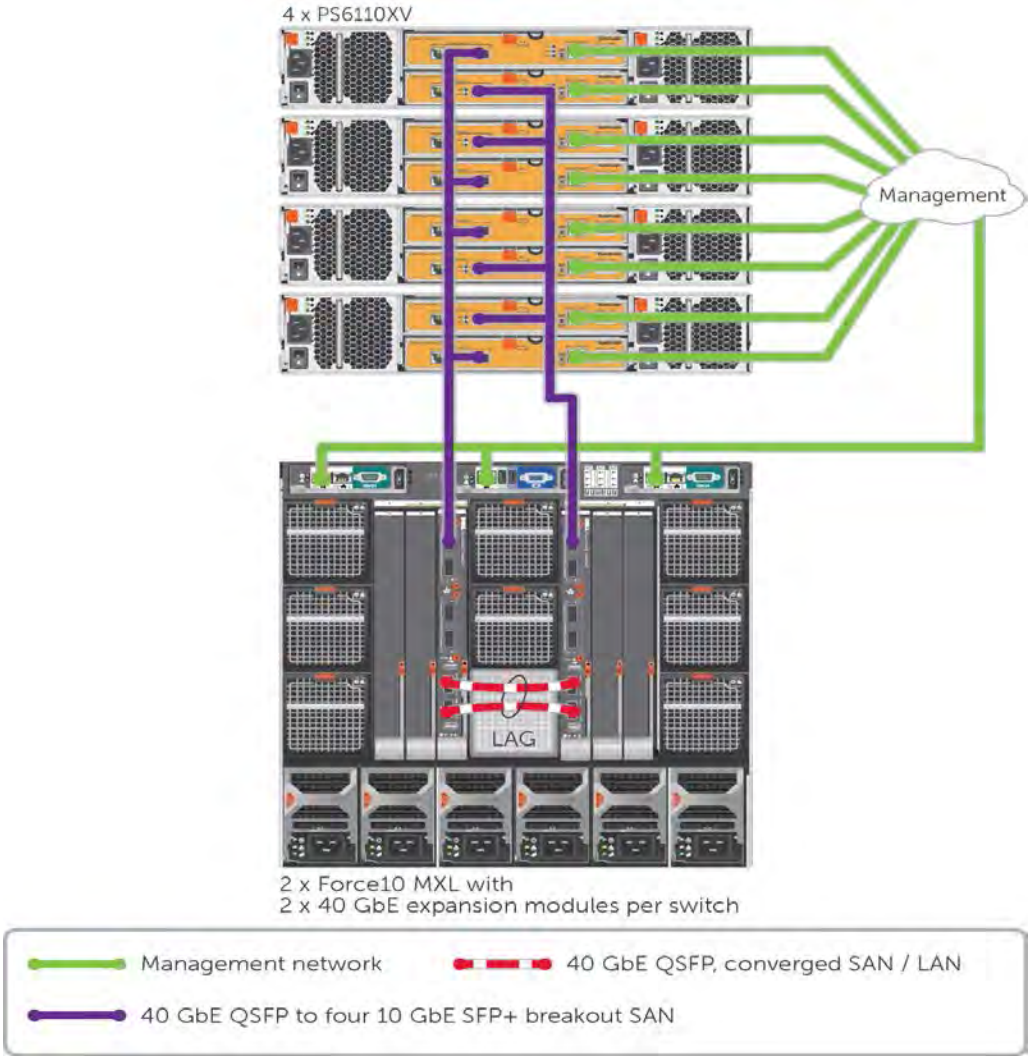

4 x PS6110XV
Management
LAG
2 x Force10 MXL with
2 x 40 GbE expansion modules per switch
Management network
40 GbE QSFP, converged SAN / LAN
40 GbE QSFP to four 10 GbE SFP+ breakout SAN

Figure 48 Blade IOM switch only

14.5.2 ToR switch only

Network ports of both the hosts and the storage are connected to external ToR switches. 10 GbE pass-through IOM switches are used in place of blade IOM switches in the M1000e blade chassis. The switch interconnect can be a stack, a LAG, or a VLTi.

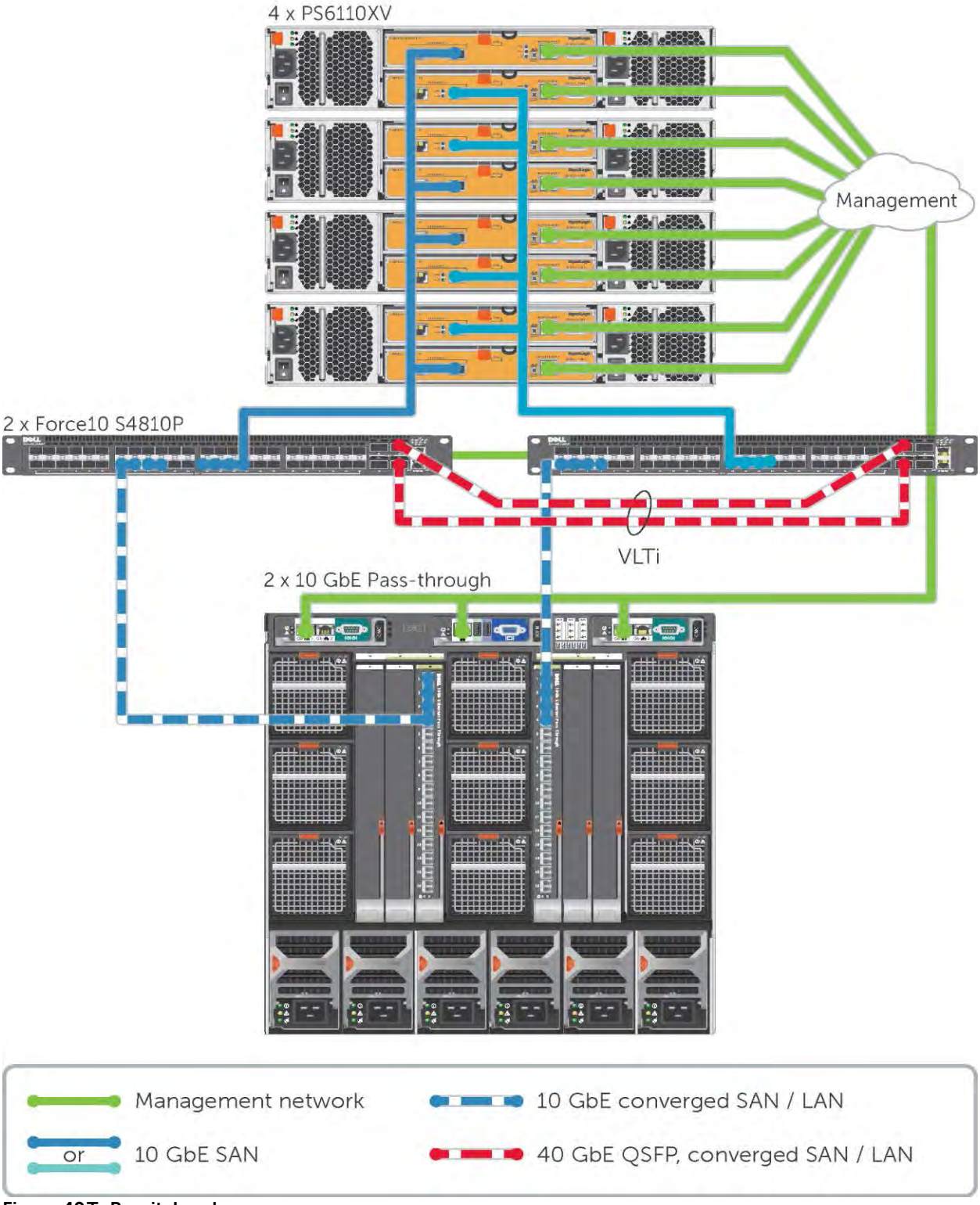


Figure 49ToR switch only

14.5.3 Blade IOM switch with ToR switch

Host network ports are connected to the M1000e blade IOM switches and the storage network ports are connected to ToR switches. The switch interconnect can be a stack, a LAG, or a VLTi and should

connect the ToR switch to better facilitate inter-array member traffic. An uplink stack, LAG, or VLT LAG from the blade IOM switch tier to the ToR switch tier is also required.

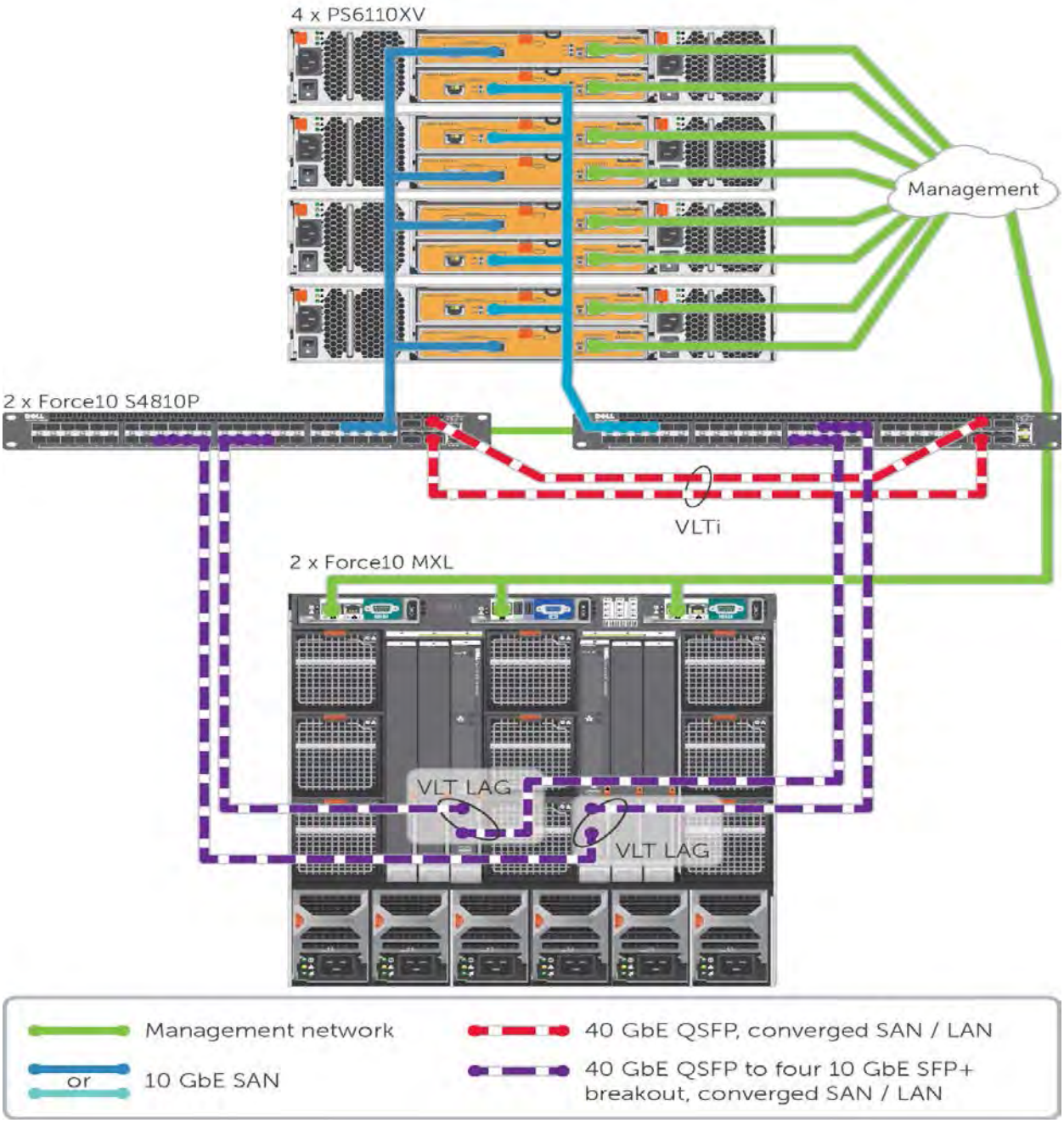


Figure 50 Blade IOM switch with ToR switch (Two tier design)

14.6 Data Center In A Chassis DCB design

Find guidance for incorporating a data center-in-a-chassis with the new Dell EqualLogic PS-M4110 Blade Array, Force10 Blade IO Modules, and PowerEdge M-Series blade servers to form a reliable, stable, and well performing full end-to-end DCB solution.

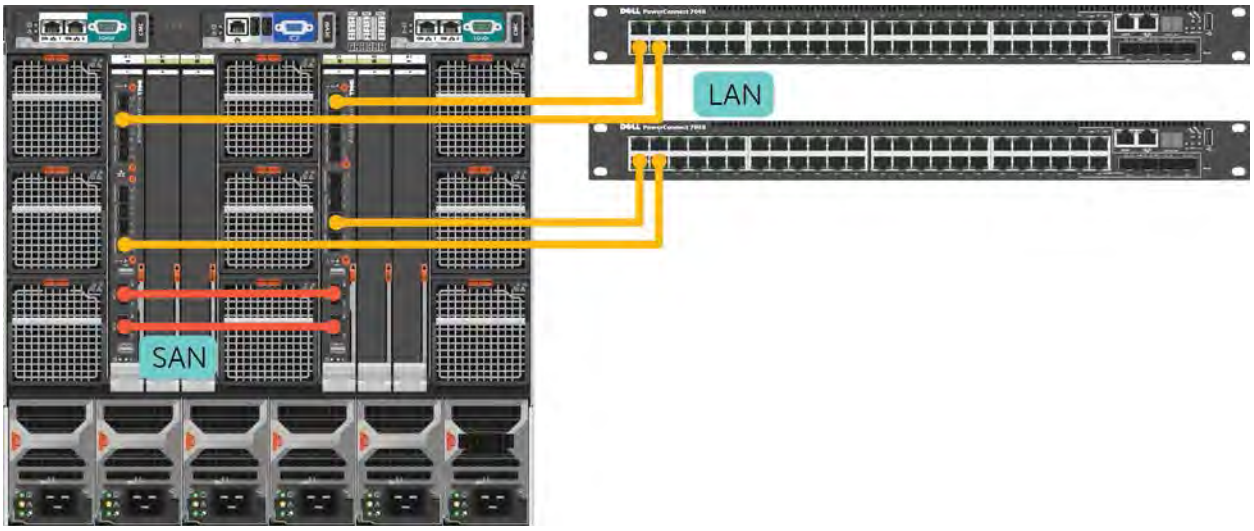


Figure 51 M1000e Blade Enclosure full DCB solution

The components used in the solution are listed here:

- M1000e chassis
- M620 blade server
- PS-M4110XV array
- Force10 MXL (2)
- Broadcom 57810 CNA

14.7 VLANs for iSCSI

A non-default VLAN is required for operating prioritized lossless iSCSI traffic in a DCB enabled Ethernet infrastructure. Switch ports that are based on the IEEE 802.1Q VLAN specification forward frames in the default or native VLAN without tags (untagged frames). Normally these ports also receive frames in the default or native VLAN as untagged. Such a port is typically termed a "tagged port" or "trunk port" and all non-default VLAN frames are forwarded with the tags intact. Since the DCB priority information (PCP value) is encoded in the VLAN tag, this information will be lost if an end-device sends iSCSI traffic in a default or native VLAN and the switch receives it in a tagged or trunk port. As a result, DCB prioritization and traffic classification will be lost.

All devices in the iSCSI data path must have the same VLAN ID configured on the respective ports participating in the iSCSI network to ensure proper functioning. These devices include the server iSCSI NIC/CNA ports, EqualLogic arrays, and all switches on the iSCSI SAN.

> **Note:** The VLAN ID for iSCSI can be set in the EqualLogic Group Manager interface or the storage array CLI. When DCB is enabled on the switch, it is necessary to configure a non-default VLAN on the array, switch, and all host ports that are part of the EqualLogic SAN. VLAN IDs 0 and 1 are not recommended, because these may be the default or reserved VLAN for some switches, and as such, may forward frames untagged (e.g. no VLAN tagging). VLAN tagging is required to fully support DCB.

For more information on the DCB requirements and configuration guidelines, see the following white paper: Data Center Bridging: Standards, Behavioral Requirements, and Configuration Guidelines with Dell EqualLogic iSCSI SANs

For more information on the DCB requirements for EqualLogic, to ensure that DCB is properly enabled and/or disabled across all devices, and to assist with identifying and resolving basic DCB configuration issues in the EqualLogic SAN, see the following white paper: EqualLogic DCB Configuration Best Practices

For a comprehensive understanding of integrating an M1000e Blade Chassis in a Full End-to-End Data Center Bridging environment, see the following white paper: M1000e Blade Enclosure and EqualLogic Arrays SAN Design using Force10 Switches

Appendix A Network ports and protocols

PS Series groups use a number of TCP and UDP protocols for group management, I/O operations, and internal communication. If you have switches or routers set to block these protocols, you may need to unblock them to allow management or I/O operations to work correctly. The required and optional protocols are listed in the following sections.

A.1 Required ports and protocols

Table 49 lists the ports and protocols required for operating an EqualLogic iSCSI SAN.

Table 49 Required ports and protocols

Type	Port	Protocol	Access
iSCSI			
TCP	3260	iSCSI	To the group IP address and all individual member IP addresses
EqualLogic Internal			
UDP	161	SNMP	Management operations
TCP	9876	Internal	iSCSI intra-system control
TCP	25555	Internal	Group communication
TCP	20002	Internal	Event logging

A.2 Optional ports and protocols

Table 50 lists the optional ports and protocols used for management and alerts. They are not required for correct array operation.

Table 50 Optional ports and protocols

Type	Port	Protocol	Access
CLI Management			
TCP	23	Telnet	To group IP address
TCP	22	SSH	To group IP address
Web Based Management			
TCP	80	HTTP	To group IP address
TCP	3002	GUI communication	To group IP address
TCP	3003	GUI communication (encrypted)	To group IP address
SNMP			
UDP	161	SNMP	To and from group IP address
Syslog			
UDP	514	Syslog	From group IP address
EqualLogic Diagnostics			
TCP	20	FTP	Software update and diagnostic procedures; to all individual member IP addresses
TCP	25	SMTP	E-mail and diagnostic notifications; from all individual member IP addresses to the configured SMTP server

Appendix B Recommended switches

The list of recommended switches is now maintained in a separate document.

- EqualLogic Compatibility Matrix
  http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19856862/download.aspx

Appendix C Supported iSCSI initiators

The list of supported iSCSI initiators is now maintained in a separate document.

- EqualLogic Compatibility Matrix
  http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19856862/download.aspx

Appendix D Upgrade paths for EqualLogic PS SeriesArrays

Table 51 EqualLogic Upgrade Paths

End of sales life arrays:	Latest available conversion model	1Gb to 10Gb conversion availability	Drive upgrades availability
PS-50 thru PS3000	None	None	None
PS4000	Yes, PS6000 – Dellstar	Yes, PS6010 – Dellstar	Yes – cus kit tool
PS5000	Yes, PS6000 – Dellstar	Yes, PS6010 – Dellstar	Yes – cus kit tool
PS6000	None	Yes, PS6010 – Dellstar	Yes – cus kit tool
PS5500	Yes, PS6500 – Dellstar	Yes, PS6510 – Dellstar	Yes – cus kit tool
PS6010	None		Yes – cus kit tool

Currently shipping arrays:	Latest available conversion model	1Gb to 10Gb conversion availability	Drive upgrades availability
PS6500	None	Yes, PS6510 – Dellstar	Yes – cus kit tool
PS6510	None		Yes – cus kit tool
PS4100	TBD	TBD	Yes – cus kit tool
PS4110	TBD		Yes – cus kit tool
PS6100	None	TBD	Yes – cus kit tool
PS6110	None		Yes – cus kit tool
PS-M4110	None		Yes – cus kit tool

Related publications

The following locations provide additional background and technical details supporting configuration of EqualLogic SANs.

- EqualLogic Compatibility Matrix
  http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19856862/download.aspx
- EqualLogic Technical Content
  http://en.community.dell.com/techcenter/storage/w/wiki/2660.equallogic-technical-content.aspx
- Rapid EqualLogic Configuration Portal
  http://en.community.dell.com/techcenter/storage/w/wiki/3615.rapid-equallogic-configuration-portal-by-sis.aspx
- Switch Configuration Guides
  http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-by-sis.aspx
- Storage Infrastructure and Solutions Team Publications
  http://en.community.dell.com/techcenter/storage/w/wiki/2632.storage-infrastructure-and-solutions-team-publications.aspx

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.